Journal of
Clinical Psychology

JULY 1960
Volume XVI Number 3

EDITORIAL BOARD

FREDERICK C. THORNE, *Editor*
Editorial Office: 5 Pearl Street
Brandon, Vermont

JERRY W. CARTER, JR.
U. S. Public Health Service

HOWARD F. HUNT
University of Chicago

WILLIAM A. HUNT
Northwestern University

ELAINE F. KINDER
One Haven Court
Nyack, N. Y.

MAURICE LORR
Veterans Administration
Washington, D. C.

BERNARD SAPER
N. Y. State Department of Mental Hygiene

SAUL B. SELLS
Texas Christian University

ROBERT I. WATSON
Northwestern University

ELEANOR C. THORNE, *Business Manager*

Journal of Clinical Psychology
5 Pearl Street · · Brandon, Vermont

COMPANION VOLUMES

PRINCIPLES OF PSYCHOLOGICAL EXAMINING

by

FREDERICK C. THORNE, M.D., Ph.D.
Editor, Journal of Clinical Psychology

PRICE: $7.50

PRINCIPLES OF PERSONALITY COUNSELING

by

FREDERICK C. THORNE, M.D., Ph.D.
Editor, Journal of Clinical Psychology

PRICE: $6.00

THE HOWARD INK BLOT TEST

by JAMES W. HOWARD, Ph.D.

NEW - - - PRICE: $5.00 - - - ORDER NOW

MENTAL DEFICIENCY

by J. E. WALLACE WALLIN, Ph.D.

PRICE: $5.00 - - - ORDER NOW

CLINICAL PSYCHOLOGY MONOGRAPHS

1. Billingslea: Bender Gestalt Test.....$ .90
2. Carter & Bowles: Qualitative Aspects of Psychological Examining.....$1.25
3. Training in Clinical Psychology (Available only in Complete Sets).....
4. Critique of Nondirective Psychotherapy.....$1.25
5. Buck: The House-Tree-Person Test Manual.....$2.00
6. Karpman: Objective Psychotherapy.....$2.00
7. Statistics for the Clinician: A Symposium.....$2.00
8. Tyson: Current Mental Hygiene Practices.....$2.00
9. Counseling with the Mentally Defective.....$1.25
10. Howard Ink Blot Test Manual.....$2.00
11. Ellis, Albert: New Approaches to Psychotherapy Techniques.....$2.00

NEW

12. Gardner Lindzey, Jean Bradford, Charlotte Tejessy and Anthony Davids
THE THEMATIC APPERCEPTION TEST: An Interpretive Lexicon.....$2.50
All TAT workers will wish to refer to this authorative compendium of interpretations summarized from the research literature.
13. Dell Lebo: GUIDEBOOK FOR THESIS COMPOSITION.....$1.25
This monograph summarizes the mechanics of thesis writing. Invaluable as an outline for graduate students.

ORDER NOW! FROM

JOURNAL OF CLINICAL PSYCHOLOGY

5 Pearl Street Brandon, Vermont

JOURNAL OF CLINICAL PSYCHOLOGY

VOL. XVI | JULY 1960 | No. 3

CONTENTS

STUDIES OF PERSONALITY DEVELOPMENT AND STRUCTURE

The Objective Measurement of Motivation Structure in Children. RAYMOND B. CATTELL, ARTHUR B. SWENEY and JOHN A. RADCLIFFE - - - - 227

Need-Attitudes of Normal and Paranoid Schizophrenic Males. JAY L. CHAMBERS and LOUIS J. BROUSSARD - - - - - - - - - - - 233

Psychosexual Development in Terms of Object and Role Preferences. VLADIMIR PISHKIN - - - - - - - - - - - - - - - - - 238

PSYCHIATRIC RATING SCALES AND MEASURES

The Psychotic Reaction Profile. MAURICE LORR, JAMES P. O'CONNOR and JOHN W. STAFFORD - - - - - - - - - - - - - - - - 241

A Ward Behavior Rating Scale for Mental Hospital Patients. E. I. BURDOCK, G. HAKEREM, ANNE S. HARDESTY AND JOSEPH ZUBIN - - - 246

Reliability-Unreliability of Ancillary Psychiatric Evaluations. MILTON B. JENSEN and WILLIAM E. MORRIS - - - - - - - - - - - - 248

The Relation of Person Naming and Tension to Prognosis and Adjustment of Chronic Mental Patients. JEFFREY CAINE and GERDA RAND FILLENBAUM 252

The Attitude of Industrial Employers toward Hiring of Former State Mental Hospital Patients. VYTAUTAS J. BIELIAUSKAS and HARVEY E. WOLFE - 256

Attitudes of Released Chronic Schizophrenic Patients Concerning Illness and Recovery as Revealed by a Structured Post-Hospital Interview. G. W. BROOKS and W. N. DEANE - - - - - - - - - - - - - - - 259

The Bender-Gestalt Test as a Predictor of Length of Hospitalization with Mental Patients. DALE S. HIGBEE, JOHN R. CLARK and WAYNE E. HENDERSON - 265

STUDIES OF VISUAL AND MOTOR FUNCTIONS

Speed and Quality of Psychomotor Performance in Psychiatric Patients. M. B. SHAPIRO, R. KESSELL and A. E. MAXWELL - - - - - - - 266

Distance Constancy in Schizophrenic and Non-schizophrenic Patients. T. E. WECKOWICZ and R. HALL - - - - - - - - - - - - - - - 272

A Comparison of the Bender-Gestalt and Flicker Fusion as Indicators of Central Nervous System Involvement. FREDERICK L. MCGUIRE - - 276

Comparison of Reproduction and Recall Size Deviations in the Bender-Gestalt as Measures of Anxiety. DANIEL GAVALES and THEODORE MILLON - 278

Children's Spontaneous Object Elaborations on a Visual-Motor Test. G. H. ZUK - 280

MINNESOTA MULTIPHASIC PERSONALITY INVENTORY STUDIES

A Factor Analytic Study of Barron's Ego Strength Scale. EVELYN CRUMPTON, JOEL M. CANTOR and CURT BATISTE - - - - - - - - - - 283

The Representation of MMPI Scales by MMPI Factor Scales. EDWARD F. GOCKA and HAYDEN L. MEES - - - - - - - - - - - - - - 291

Two MMPI Scales for Alcoholism: What Do They Measure? MAURICE KORMAN - 296

CONTENTS—Continued

MMPI High Point Code Frequencies in a State Hospital Population. REUBEN J. SILVER and LLOYD K. SINES - - - - - - - - - - 298
A Shortened Schizophrenic Scale: Application to Confined Groups. RALPH ROBINOWITZ - - - - - - - - - - - - - - - - - - 301
MMPI Profiles of Psychopaths and Prisoners. LEONARD KINGSLEY - - 302
Comparison of MMPI Profile Differences among Negro-White Adolescents. JOHN C. BALL - - - - - - - - - - - - - - - - - - 304
MMPI Performance of Aphasic and Nonaphasic Brain-Damaged Patients. DONALD G. DOEHRING and RALPH M. REITAN - - - - - - - - - 307
Two Types of Paranoid Schizophrenia. BENJAMIN KLEINMUNTZ - - - 310

STUDIES OF WECHSLER SCALES
Beta-WAIS Comparisons and WAIS Subtest Configurations within a State Prison Population. JAMES H. PANTON - - - - - - - - - - - 312
The Validity of WAIS Performance Subtests Completed with One Hand. PETER F. BRIGGS - - - - - - - - - - - - - - - - - - 318
The Performance-Verbal IQ Discrepancy in a Group of Sociopaths. J. G. FIELD - 321
The Relationship between the CMAS and WISC Functioning. A. JACK HAFNER, DONALD M. POLLIE and IRWIN WAPNER - - - - - - - 322

STUDIES OF PROJECTIVE MEASUREMENTS
Rorschach Indices and Autonomic Reactivity. J. E. TONG and I. C. MURPHY 324
The Projective Effects of Fear and Sexual Arousal on the Rating of Pictures. ROBERT H. HORNBERGER - - - - - - - - - - - - - - - - 328
The Incomplete Sentences Blank. Sources of Variance in Retest Reliability. MARK W. STEPHENS - - - - - - - - - - - - - - - - - - 331
Differences in Performance on the Chromatic vs. Achromatic H-T-P Drawings. VYTAUTAS J. BIELIAUSKAS and ANN R. HEFFRON - - - - - - 334

MISCELLANEOUS PAPERS
Interrelationships among Porteus Maze Test Qualitative Errors. RICHARD F. DOCTOR - - - - - - - - - - - - - - - - - - - 336
AVA and the Kessler PD Scale as Measures of Passive-Dependency. PETER F. MERENDA, WALTER V. CLARKE and SYDNEY KESSLER - - - - - 338
The Effects of Two Simultaneous Cognitive and Affective Stimuli on a Group of Chronic Schizophrenic Patients with Suicidal Ideation. EMANUEL STARER - 341
Operational Psychological Report Writing. FREDERICK C. THORNE - - 343
The Utility of the Progressive Matrices (1956 Edition) with Normal and Retarded Children. LESLIE F. MALPASS, RONALD BROWN and DONALD HAKE - 350
A Comparison of WAIS Scores in Hemiplegic Patients With and Without Aphasia. SAMUEL HIRSCHENFANG - - - - - - - - - - - - 351
Effect of Mental Illness on Intelligence Test Scores. RICHARD M. GRIFFITH and VIVIAN H. TAYLOR - - - - - - - - - - - - - - 352

Editorial Opinions - - - - - - - - - - - - - - - - - - 353

GENERAL INFORMATION

The *Journal of Clinical Psychology* is published quarterly, appearing in January, April, July and October, at Rutland, Vermont, by Frederick C. Thorne, M.D. Subscription price $10.00; Canadian subscriptions $10.50; foreign subscriptions $11.00 including postage; single copies $2.50. Entered as second class matter January 15, 1945, at the post office at Burlington, Vermont, under the Act of March 3, 1879. Reentered at the post office at Brandon, Vermont. Additional entry at the post-office at Rutland, Vermont. *Title registered, U. S. Patent Office.* Editorial and Business Offices, 5 Pearl Street, Brandon, Vermont.

Copyright, 1960, Frederick C. Thorne, M. D.

THE OBJECTIVE MEASUREMENT OF MOTIVATION STRUCTURE IN CHILDREN[1]

RAYMOND B. CATTELL AND ARTHUR B. SWENEY
University of Illinois

JOHN A. RADCLIFFE
University of Sydney

Problem

The application of a number of objective tests to determine motivation components is opening a new dimension in the study of attitudes. Cattell and Baggaley [6, 7] studying the attitudes of Air Force men isolated five replicable and psychologically meaningful factors. Three were descriptively identified with psychoanalytic functions ascribed to the id, ego and super ego, and the other two factors were descriptively related to the physiological involvement in an attitude and the repressed conflicts which an attitude summons forth. A previous study by Radcliffe [8] with a limited number of variables yielded results which largely confirmed the presence of these same motivation components in the measurement of children's attitudes. That study, like Tapp's [10] utilized group devices and therefore failed to isolate a delta factor.

The purpose of this study was to replicate the three previous studies mentioned and to add instruments to measure the physiological components. It assumed the additional function of testing new devices utilizing psychodynamic principles which might clarify the manifest nature of the motivation factors thus far only descriptively identified.

Method

Subject. The subjects for this study, as for the pilot study, were grade school children of both sexes from the public schools of Decatur, Illinois. The racial constituency was largely white, but included a few Negroes of insufficient number to warrant statistical recognition.

Attitudes. To conform to the general pattern of the pilot studies [8, 10] this study was also primarily directed toward the measurement of attitudes toward movies and religion. Attitudes from eight other sentiment areas were included to be used as an "attitude pool". These other areas were: sports and fitness, comics, patriotism, science, domestic tasks, out-of-doors, mechanical activities and pets. The religious attitude to be measured was "I want to go to church to gain help from my faith". The attitude toward movies was "I want to go to the movies".

Test Devices. The test devices were carefully selected to fill several needs distinctive to this study. The criteria utilized in this choice were: historical effectiveness as markers of motivation components, adaptability physically and psychologically to a young population, and ability to discriminate between two or more conflicting hypotheses for the identifying of the same motivation components. Marker variables were chosen largely from the Cattell and Baggaley [6, 7] studies.

Statistical Analysis. The scoring of the test devices was largely ipsative, thus measuring not differences between individuals but differences between any individual's attitudes. These test scores for religion and for movies were intercorrelated into two separate matrices, each of which was then factored. Preliminary studies suggested that the extraction of fifteen factors should be adequate to include the significant interrelationships.

The centroid factor structure was then subjected to standard oblimax rotation process as programmed on the Illiac computer. This, essentially, is a process of finding an oblique position for each reference vector which maximized the fourth power of the projections upon it by the variables. A further visual rotation for simple struc-

[1]This research was sponsored by Research Grant #1715, Mental Health Committee, Department of Public Welfare, State of Illinois.

ture and a maximized hyperplane was carried out by over fifteen iterative steps for each of the two portions of the study.

Results

The same factor patterns found by Cattell and Baggaley were replicated in this study (see Tables 1 and 2) in varying degrees of exactness. Two new motivation factors were found which exhibited similar factor loadings for both attitude studies and were therefore distinguishable from the artifacts related to experimental conditions.

Table 1. The Oblique Vector Structure from Blind Rotation for Movies (N = 89)

	Alpha	Beta	Gamma	Delta	Epsilon	Zeta	Eta
Preference	.66	—	—	.20	—	—	—*
Choice to Explain	.52	—	—	—	—	—	—
Id Projection	.38	—	—	—	—	—	—
Visual Attention	.31	—	—	—	—	—	—
Auditory Attention	.39	—	—	—	—	—	−.23
Naive Projection	.33	—	—	—	—	—	—
True Projection	—	—	—	—	—	—	—
Autism	—	—	—	—	—	—	—
Information	—	−.20	.60	—	—	—	—
Learning Language**	—	−.56	—	—	—	—	—
Free Association Speed	—	−.56	—	—	—	—	−.25
Systolic Pressure	—	−.35	—	−.32	—	—	—
Diastolic Pressure	—	−.36	—	—	—	.27	—
Perceptual Integration	—	.37	—	−.38	—	—	—
Word Association	—	—	.50	—	—	—	—
Paired Words	—	—	.20	—	—	—	—
Selective Perception	—	—	.38	—	.26	—	—
Super-ego Projection	—	—	.28	—	—	—	—
Identification	—	—	.22	—	—	—	—
P. G. R. Deflection	—	—	—	.53	—	—	—
Decision Speed (Opinion)	—	—	—	.45	—	—	—
Memory	—	—	—	−.25	−.38	—	−.26
Reminiscence	—	—	—	—	.49	—	—
Reaction Speed (Concensus)	—	—	—	—	−.41	.71	—
Decision Strength	—	—	—	—	—	.27	—
Impulsiveness	—	—	.38	—	—	.25	—
P. G. R. Drift	—	—	—	—	—	−.25	—
Intelligence	—	—	—	—	—	—	−.71
Super-ego Strength	—	—	—	—	—	—	−.31
Maintained Fluency	—	—	—	—	—	—	.42
Maintained Perceptual Task	—	—	—	—	—	—	.60
Maintained Motor Task	—	—	—	—	—	—	—

*Dashes represent loadings below the criteria of .20.
**Applied to movies only.
***The following tests loaded none of the components sufficiently to meet the criteria: true projection, ego projection, visual resolution, and warm up time.

α *Alpha Factor: Conscious Id or Degree of Satisfaction Currently Achieved.*

The factor which was most dominant and which was matched on the greatest number of variables corresponded to the factor alpha of the Cattell and Baggaley [6, 7] studies and of the Radcliffe [8] pilot study. This motivation component has been hypothesized to be related to the id forces operating in attitude behavior. An alternative, not completely conflicting, hypothesis has defined it as "hedonic energy", while, if we attempt alignment of M. C. factors, especially those in the second-order

TABLE 2. THE OBLIQUE VECTOR STRUCTURE FROM BLIND ROTATION FOR RELIGION (N = 89)

	Alpha	Beta	Gamma	Delta	Epsilon	Zeta	Eta
Preference	.44	—	—	—	.43	—	—*
Choice to Explain	—	—	—	—	—	—	—
Id Projection	—	—	—	—	—	.34	—
Visual Attention	.42	—	—	—	—	—	
Auditory Attention	.43	.23	.23	—	—	—	—
Naive Projection	.46	—	—	—	—	—	—
True Projection	−.45	—	—	—	.23	.21	—
Autism	.29	—	—	—	.30	—	.32
Information	—	.31	.20	—	—	—	—
Distort of Reasoning**	.38	—	—	—	—	—	—
Free Association Speed	—	.44	—	—	−.23	—	—
Systolic Pressure	—	−.38	—	−.44	—	—	—
Diastolic Pressure	—	−.28	—	−.23	—	—	—
Perceptual Integration	—	—	−.37	—	—	—	—
Word Association	—	—	.49	—	—	—	—
Paired Words	.37	—	.27	—	—	—	—
Selective Perception	—	—	.37	—	—	—	—
Super-ego Projection	—	—	—	—	—	—	—
Identification	—	—	—	—	—	−.34	—
P. G. R. Deflection	−.21	—	—	.48	—	—	—
Decision Speed (Opinions)	—	—	—	.39	—	—	—
Memory	—	—	—	—	−.45	—	—
Reminiscence	—	—	—	—	.49	—	—
Decision Speed (Concensus)	—	—	—	—	—	.50	—
Decision Strength	—	—	—	—	—	.56	—
Impulsiveness	—	—	−.37	—	—	.30	—
P. G. R. Drift	−.22	—	—	—	—	—	—
Intelligence	—	—	—	—	—	−.34	−.43
Super-ego Strength	—	—	—	—	—	—	−.56
Maintained Fluency	—	—	—	—	—	—	.23
Maintained Perceptual Task	—	—	—	—	—	—	.25
Maintained Motor Task	—	—	—	—	—	—	.26

*Dashes represent loadings below the criteria of .20.
**Applied only to religion.
***The following tests loaded none of the components sufficiently to meet the criteria: ego projection, and warm up time.

"unintegrated interests" (A) factors with E terms, i.e., with terms in the ergic tension equation, we would identify it as G(−), i.e., *lack* of satisfaction after stimulation. The basis for regarding it as that component which corresponds to a deficiency of outlets for satisfaction is the high awareness of need, and the general autistic "I wish quality". If this identification is made, then this factor should show some entanglement with the η factor below which could be identified as stimulation intensity factor.

β Beta Factor: Ego Component, or Established Sentiment Factor Ingredient.

This study has replicated previous studies in the fact that, as before, the beta or ego component has been somewhat indistinct. Its most consistent variable has been information concerning the *means* for satisfying the attitude. A certain uniformity of signs and low loading variables helps identify this factor in the present study. The lack of preference, added to the integrated nature of this factor, serves to delineate its function from the alpha or id component quite successfully. This factor has the characteristic of all sentiment factors that it is strongest in *realized, habituated* interests. However, since γ similarly appears consistent with attitudes

that have become learned, or are in process of being learnt, it is not possible simply to equate β with the contributions made by the M terms in the specification equation. Instead, we are forced to conclude that each M term in some way has two components, β and γ. Most likely we shall have to conclude that *every* sentiment is involved either in the ego or super ego structure. (Possibly the latter will appear as second-order factors among sentiment measures.) Meanwhile, still consistently with the psychoanalytic conception, we shall hypothesize β as an ego component in all sentiments (M terms).

γ *Gamma Factor: Ideal Self, Super Ego or Aspired Sentiment Factor Ingredient.*

The position of gamma or super ego in the children's population, represented by this sample, was less prominent than alpha but easier to identify than beta. In this case, on the integrated level, word association and information loaded consistently. The tests of distortion of reasoning and choice to explain were loaded in the proper direction but less heavily than in either the pilot study or the previous adult studies.

The wishful nature of gamma is emphasized by autism and distortion of reasoning. The super ego projection test is constructed to measure rationalization for possessing an attitude. The former "super ego" interpretation is retainable, but still more appropriate is the conception that this is an ideal self component strong in every newly developing sentiment or M term. Unfortunately for increased confidence in this view, the identification test failed to load this factor as it was expected.

δ *Delta Factor: Physiological Need Component.*

This factor has been hypothesized to be the physiological component of motivation (P in the ergic tension equation). Its loading pattern in this study follows quite faithfully those found in the adult study of Cattell and Baggaley. The magnitudes of the loadings are noticeably smaller, but within the range expected. The high decision speed and high deflection of the P. G. R. to threat seem to pin point this factor as the component related to skeletal and autonomic facilitation. The other variables in the factor seem consistent with this hypothesis.

ϵ *Epsilon Factor: Motivation Tied to Unconscious Memories (Arising From Complexes).*

The nature of epsilon is under some scrutiny. It was been hypothesized as the component representing the effects of repressed complexes. This study paralleled the one by Cattell and Baggaley by indentifying this factor by high but reversed loadings on memory and reminiscence. The P. G. R. variables in this study were only negligibly found in this factor. The hyperplane or negligibly loaded variables, however, are consistent both in sign and magnitude to the expected results.

ζ *Zeta Factor: Constitutional Need or Urgency.*

One of the factors which seems to be unique in this study is the one which may be characterized as an urgency and impulsiveness in the quality of interest. This component is highly loaded by quickness of decision (saying "yes") upon commonly agreed upon attitudes, and in the test of impulsiveness, which measures the tendency to select an early answer from a temporal series of equivalent answers to an attitude question. It is also substantially loaded on decision strength which is a measure of the absolute displacement of a preference lever against a resisting spring as the subject registers an affirmative or negative answer to an attitude question, i.e., it measures the forcefulness of motor expression with which the interest is expressed.

η *Eta Factor:*

The control characteristic of this factor, which has not previously appeared clearly, is seen in loadings on fluency (second minute, not first), maintaining activity in skeletal words, and ego projection. Also, as far as movies and religion are concerned, this kind of activity is found more in persons of low intelligence, low superego strength and high surgency. Is it possible that we are dealing with a high stimulation

or low satisfaction phenomenon, which, in persons of lower intelligence, is particularly likely to occur in movies and religion?

In terms of direct description of behavior, η might be designated a tendency "to go on and on with not-too-well-directed mental activity about a topic," which is not incompatible with the notion that the person is under more stimulation of the ergs normally getting satisfaction through those channels than he can discharge. The fact that verbal, conscious preference has no marked loading is compatible if we consider the stimulation recent so that the individual has had insufficient opportunity to recognize his needs.

The remaining factors visible in the Tables are essentially doublets from a single test or obvious artifacts arising from the design of this study or to the nature of the test variables utilized. They will therefore not be listed here but hypotheses as to the nature of their occurrence will be developed in the discussion section.

Test Variables. Table 2 exposits the characteristics and loadings of the test variables utilized in this study. Further information about individual test devices in relation to their performance in measuring primary motivation strength in children was one of the primary aims of the study. For this reason their characteristics become important for the evaluation of this study and for the development of future ones.

Discussion

The relationship of the motivational components found in this study to the characteristics posited upon the id, ego and super ego makes it possible to discuss the nature of the factors in light of psycho-social development. This therefore becomes a process of extending and expanding the inferential sub-structure of the measured factor.

Alpha or the id component is by far the most heavily and consistently loaded factor. Although selection of test devices may enter into this, it may be noted that the resistance to auditory distraction which was related to beta in the adult study is very consistently related to alpha in this study. The prevalence of the id component in children's attitudes would certainly not be inconsistent with psychoanalytic theory.

The difficulty of isolating Beta also may be related to psychosexual development. Most sources agree that the ego emerges historically from the conflict between id drives and external restrictions, but expresses itself only in an expanded field of choices. Thus one would expect the Beta component to be somewhat smaller in the child than in the adult. The alternative interpretation of the Beta factor as an integrated component of motivation would also not be inconsistent with its depression in the early years. The process of conscious integration of methods of drive satisfaction is certainly related to cultural assimilation and thereby to age.

The identification of the Gamma component with the one found in adults hinged very largely upon word association and choice to explain. The strong presence of information in this factor may be significant developmentally. Assuming that the super ego does result from incorporation of external requirements and restrictions, this factor would suggest that the child's super ego is still largely conscious. For this reason, information concerning his attitudes or motivation is still important for satisfaction within the restrictions of society. It would therefore be expected that gamma and beta would be more closely related in the child than the adult, and that therefore this would be reflected in the second-order factor structure.

The Delta factor is possibly the least controversial of the motivation components found. It seems significant that the inclusion of decision time and strength makes it possible to consider this factor as being both skeletal and autonomic in nature. Since there are not other variables which systematically load this factor, it seems very likely that it will not lend itself to identification via group testing and will probably therefore have to be omitted from much less intensive research.

The Epsilon factor which has been equated to effects of incompletely repressed complexes could be also interpreted as aberration of memory functions systematically related to attitudes. In past studies where the Zeigarnic effect was measured, this later interpretation would certainly be consistent. In this study lack of other consistent loadings makes interpretation difficult. The fact that information and fluency appear in the hyperplane would suggest the unintegrated quality of this factor.

Conclusions

This study applied a large number of different motivational devices to the study of a limited number of attitudes in children. It was found that the motivational factor structure in children was highly consistent with the structure found by Cattell and Baggaley in their study of adults. Distinct factors could be descriptively identified with the psycho-analytical id, ego and super ego. The presence of a physiological factor and one which can be related to repressed complexes also replicates the adult pattern.

Two new factors were found which were not present in any of the previous studies. One of these factors is associated with a selective readiness or impulsiveness in attitude behavior. The other factor seems to be related to fatigue and symbolic satisfaction of drives, and therefore takes the form of differential persistence. This might be also defined as differential pervasiveness of attitude in the face of obstruction.

The same general motivation patterns were found for the widely divergent attitudes toward movies and religion, using a pool of eight other attitudes for a control. From this it is concluded that the motivation components found have a certain general application.

Bibliography

1. Cattell, R. B. The ergic theory of attitude and sentiment measurement. *Educ. Psychol. Measmt.*, 1947, *7*, 221-246.
2. Cattell, R. B. Personality and motivation theory based on structural measurement. In. J. L. McCary (Ed.), *Psychology of personality.* New York: Logos, 1956.
3. Cattell, R. B. *Personality and motivation structure and measurement.* New York: World Book Co., 1957.
4. Cattell, R. B. The dynamic calculus: a system of concepts derived from objective motivation measurement In G. Lindzey (Ed.). *The assessment of human motives.* New York: Rinehart, 1958.
5. Cattell, R. B. The dynamic calculus: concepts and crucial experiments. Nebraska Symposium: University of Nebraska Press, 1959.
6. Cattell, R. B. and Baggaley, A. R. The objective measurement of attitude motivation: development and evaluation of principles and devices. *J. Pers.*, 1956, *24*, 401-432.
7. Cattell, R. B. and Baggaley, A. R. A confirmation of ergic and engram structures in attitudes objectively measured. *Austral. J. Psychol.*, 1958, *10*, 287-318.
8. Cattell, R. B., Radcliffe, J. and Sweney, A. B. The nature and measurement of components in human motivation. In press.
9. Sweney, A. B. The objective measurement of motivation in children. A paper presented at the Convention of American Psychological Association, 1959.
10. Tapp, J. T. An examination of hypotheses concerning motivation components of attitude strength. Unpublished masters thesis, Univer. of Illinois, 1958.

NEED-ATTITUDES OF NORMAL AND PARANOID SCHIZOPHRENIC MALES[1]

JAY L. CHAMBERS AND LOUIS J. BROUSSARD

Charles L. Mix Memorial Fund, Inc.
Americus, Georgia

PROBLEM

The way a person feels about experienced needs may be one of the more important aspects of his motivation and adjustment. Peak[3] indicated that an evaluative attitude toward a need (*e.g.*, sex is "good" or "bad") may control or select the behavior generated by that need by exerting a governing effect on perception or determining the perspective from which the world is viewed.

Need-attitudes could be related to personality conflicts in various ways. Murray[2] suggests that attitudes toward needs might interact either to complement or conflict with each other. An attitude could also cause frustration if it prohibited a strongly felt need. Since attitudes operate at an ideational level, need-attitude conflicts would involve ideas which could become sufficiently persistent and generalized to be a factor in emotional disorders and mental illness.

Since Hovland[1] has shown that some attitudes may be altered, knowledge of pathological need-attitudes might enable a therapist to resolve some of the conflicts of his clients by modifying disturbing attitudes. This study was designed to examine some aspects of the role of need-attitudes in adjustment. Specifically, a group of normals and a group of paranoid schizophrenics were compared for differences and similarities in their attitudes toward the needs described in Murray's system.

METHOD

Subjects. Four groups of 50 S's were tested. Two groups (P1 and P2) consisted of institutionalized male patients selected from five mental hospitals diagnosed as paranoid schizophrenics by hospital staffs. The majority were newly admitted patients but ranged from first admission cases to chronic cases with up to 15 years hospitalization. The other two groups (N1 and N2) were made up of adult males with no history of mental disorder or severe emotional disturbance chosen from urban, small town, and rural communities.

The average age of the combined paranoid schizophrenic groups was 38, SD 10.3 years. The average age of the combined normal groups was 34, SD 10.8 years. The mean education level of the combined paranoid schizophrenic groups was 11.5, SD 2.75 years. The mean education level of the combined normal groups was 12.4, SD 2.31 years. Despite efforts to match groups as nearly as possible for both age and education, the combined normal and the combined schizophrenic groups differed significantly ($t > .01$) on both age and education.

PICTURE IDENTIFICATION TEST

Administration. Ten sets of eight pictures each were presented to S. Each picture was of a college boy 18 to 22 years of age, originally taken from a college annual and showed the head and shoulders of boys dressed in white shirts and ties. From each set, S was asked to select the picture of the person he would like best and the one he would like least. After S's selections were made, the first pair of pictures chosen as the person liked best and person liked least was placed before S. S was then handed a set of 21 small cards on each of which was printed a statement selected from Murray's[2] personality questionnaire designed to measure 21 needs. S was instructed to read each statement, decide which of the two depicted people it fitted best, and to place the card with the appropriate picture. The procedure was repeated until statements had been sorted for each of the ten pairs of most and least liked pictures.

[1]The Psychology Sections of the Lexington VA Hospital, the Atlanta VA Hospital, Eastern State Hospital, Virginia, Eastern State Hospital, Kentucky, and Milledgeville State Hospital, Georgia, helped in the collection of data. The assistance of Dr. Charles Ward, Dr. Richard Griffith and Mr. Frank Sgro helped make the study possible.

Scoring. If S placed a need statement with the picture of the person he liked best, he was given a + score for that need-attitude; if he placed it with the picture liked least, he was scored −. The summed score for each need-attitude was the number of +'s minus the number of −'s. Since there were ten sorts, the possible summed scores were all the even numbers and zero from −10 to +10.

If S had sorted statements in a chance or random fashion, there would be as many + scores as − scores for each need. The summed score would be zero. Therefore, the sum of the 21 summed scores, disregarding sign, indicated how much S's test result differed from chance sorting. This score was called "Variability".

If S tended to sort statements more with people liked best than with those liked least, the total of his 21 summed scores would be more positive than negative. This tendency to sort more consistently with either liked best or liked least pictures was indicated by the sum of the 21 summed scores added with regard to sign. This score was called "Consistency".

S's summed scores for each of the 21 needs were ordered from his highest positive through his highest negative score. This rank order score for each need-attitude was used as a measure of the positiveness or negativeness of S toward that need. A rank of 1 indicated the highest positive need-attitude score.

Item selection. For groups N 1 and P 1, each of the 21 needs was represented on the test by ten different descriptive statements drawn from the Murray questionnaire[2]. After the results had been obtained on these two groups, an item analysis was made of the test results of 100 adult males. The analysis indicated that not all of the ten statements gave a result as consistent with the summed score of the ten as would be optimally desirable. The five statements which agreed best with the summed score of the ten were selected for a revised test (Form II). On this alternate form, the five selected statements were used with the first five sets of pictures used on Form I and duplicates of the five selected statements were used again with the second five sets of pictures. Form II consequently provided the same number of responses as Form I on more reliable items.

Results

Comparison of Forms I and II. If Form I and Form II were comparable measures, a substantial relation would be found between the results for groups N 1 and N 2, and between the results for groups P 1 and P 2. Table 1 lists the average need-attitude

Table 1. Mean Ranks of Need-Attitudes Among Normals and Paranoid Schizophrenics

Need-Attitude	N 1 (N = 50)	N 2 (N = 50)	P 1 (N = 50)	P 2 (N = 50)
Abasement	11.5	12.6	12.2	12.1
Achievement	9.0	6.9	7.4	8.0
Affiliation	6.6	6.9	9.5	8.1
Aggression	13.1	13.0	14.0	13.5
Autonomy	12.2	13.0	12.8	13.1
Blamavoidance	10.4	11.5	10.2	9.4
Counteraction	10.0	11.3	11.2	12.2
Deference	7.8	9.9	7.9	7.9
Defendance	13.4	14.5	13.8	14.0
Dominance	10.3	8.8	9.8	10.0
Exhibition	11.0	10.5	11.5	11.9
Harmavoidance	16.5	14.0	13.7	12.8
Inferiorityavoidance	14.9	14.5	13.5	13.2
Nurturance	5.7	7.3	7.2	8.0
Order	9.2	10.0	8.8	9.2
Play	10.4	9.9	10.5	11.3
Rejection	12.4	13.5	12.7	11.5
Sentience	11.8	12.8	11.5	12.6
Sex	12.3	8.0	12.8	10.5
Succorance	12.9	14.5	11.5	12.5
Understanding	9.3	7.5	9.0	9.3

scores for each of the four groups. A correlation (rank order) between the average scores of the two normal groups was .86 and between those of the paranoid schizophrenic groups was .91. This relationship was judged high enough to consider the two forms comparable.

Analysis of Single Need-Attitude Scores. Differences were found between normals and paranoid schizophrenics with Form I on the need-attitudes of Affiliation (t = 2.57) and Harmavoidance (t = 2.31), significant beyond the 5% level of confidence. When Form II was used with the other groups of normals and paranoid schizophrenics, differences were found between two other need-attitudes, Rejection (t = 2.03) and *Sex* (t = 2.30), significant beyond the 5% level of confidence. No significant differences were found for the Variability and Consistency scores. Thus, no significant difference found between normals and paranoid schizophrenic groups by the use of one form of the test on any one variable corroborated a corresponding difference on the alternate form.

Interrelation Between Need-Attitudes. Though the average need-attitude score of normal subjects may not differ consistently from the average score of paranoid schizophrenics on any given need-attitude, differences between normals and paranoid schizophrenics might result from interactions between need-attitudes. Intercorrelations were found for all variables, the 21 need-attitudes and Variability and Consistency scores for each of the four groups. The resulting two sets of correlations for the normals were averaged by use of the z transformation, and the same procedure was repeated for those of the two paranoid schizophrenic groups. The critical ratios were then computed between the average correlations of normals and of paranoid schizophrenics. Table 2 lists all of the need-attitude pairs which provided statistically reliable differences between the normals and the paranoid schizophrenics.

TABLE 2. CORRELATIONS BETWEEN NEED-ATTIDUTE SCORES

Need Pair	Normal (N = 100)	Par. Schiz. (N = 100)	C. R.
Ach - Aff	−.25	+.19	3.07**
Ach - Bla	+.16	−.19	2.44*
Ach - Exh	−.31	+.03	2.37*
Ach - Ord	+.48	+.11	2.58**
Aff - Dfd	−.05	−.41	2.51*
Aff - Ord	−.27	+.22	3.41**
Agg - Exh	+.37	+.09	1.95
Agg - Pla	+.43	−.03	3.20**
Agg - Rej	−.10	+.25	2.44*
Bla - Dom	−.35	+.04	2.72**
Def - Ord	+.12	+.48	2.51*
Def - Cnt	−.07	+.21	1.95
Dfd - Inf	−.14	+.19	2.30*
Dom - Pla	+.19	−.14	2.30*
Exh - Ord	−.43	−.10	2.30*
Exh - Und	−.54	−.21	2.30*
Har - Pla	−.25	+.15	2.79**
Har - Var	−.47	−.11	2.51*
Inf - Ord	−.01	−.29	1.95
Nur - Rej	−.40	−.07	2.30*
Ord - Und	+.63	+.35	1.95

*C. R. of 1.96 significant at the 5% level **C. R. of 2.57 significant at the 1% level

The need-attitudes which occurred most frequently in the correlations discriminating normals and paranoid schizophrenics were: Order (6 times), Achievement (4 times), Exhibition (4 times), Play (3 times), Affiliation (3 times), Aggression (3 times), Blamavoidance (twice), Deference (twice), Inferiorityavoidance (twice), Rejection (twice) and Understanding (twice).

Cluster Analysis of Need-Attitudes. The above findings indicated that a more complex configural approach to the need-attitude variables might disclose further differ-

TABLE 3. CLUSTER ANALYSIS OF NEED-ATTITUDES OF 100 NORMAL MALES

	Need-Attitude	Avg. r	Need-Attitude	Avg. r
Cluster 1	Aggression	+.40	Order	−.36
	Sex	+.36	Understanding	−.30
	Autonomy	+.34	Achievement	−.26
	Exhibition	+.32	Blamavoidance	−.26
	Play	+.30	Abasement	−.21
Cluster 2	Aggression	+.42		
	Sex	+.35	Understanding	−.36
	Autonomy	+.33	Nurturance	−.34
	Exhibition	+.26	Deference	−.29
	Defendance	+.25		
Cluster 3	Aggression	+.30	Blamavoidance	−.30
	Play	+.28	Abasement	−.26
	Dominance	+.25	Harmavoidance	−.20
	Exhibition	+.24		
Cluster 4	Variability	+.38		
	Nurturance	+.35	Rejection	−.24
	Deference	+.34	Harmavoidance	−.21
	Affiliation	+.31		
Cluster 5	Aggression	+.40		
	Autonomy	+.34	Sentience	−.26
	Sex	+.34		
	Defendance	+.28		
Cluster 6	Sex	+.26		
	Exhibition	+.25	Succorance	−.25
	Play	+.25		
	Dominance	+.20		
Cluster 7	Counteraction	+.28	Harmavoidance	−.34
	Variability	+.25	Blamavoidance	−.21
Cluster 8	Affiliation	+.28	Harmavoidance	−.22
	Deference	+.27	Inferiorityavoidance	−.22
Cluster 9	Autonomy	+.30	Blamavoidance	−.31
			Consistency	−.21

ences between normals and paranoid schizophrenics. A cluster analysis based on the average need-attitude intercorrelations of the normal groups is presented in Table 3. Table 4 presents the results of a similar cluster analysis of the correlations found with the paranoid schizophrenic groups. The clusters are so arranged that the variables positively correlated with each other appear on the same side of the cluster. The variables on the right side of the cluster are negatively correlated with those on the left side.

Comparison of Tables 3 and 4 reveals that the largest paranoid schizophrenic cluster and the largest normal cluster contain a common core of Aggression and Autonomy opposed to Order and Understanding. Normals tend to combine positively valued need-attitudes (Exhibition, Play, Sex) with Autonomy and Aggression, while paranoid schizophrenics combine negatively valued attitudes (Harmavoidance, Inferiorityavoidance, Defendance) with Autonomy and Aggression. One possible interpretation of this finding is that Autonomy and Aggression in normals is combined with confidence, playfulness and sexuality while in paranoid schizophrenics it is combined with insecurity, defensiveness and fearfulness.

The need-attitudes positively related to Order and Understanding for normals (Abasement, Achievement, Blamavoidance) may be interpreted as indicating accomplishment and self-criticism. For the paranoid schizophrenics, the need-atti-

TABLE 4. CLUSTER ANALYSIS OF NEED-ATTITUDES OF 100 PARANOID SCHIZOPHRENIC MALES

	Need-Attitude	Avg. r	Need-Attitude	Avg. r
Cluster 1	Autonomy	+.31	Deference	−.34
	Aggression	+.28	Nurturance	−.33
	Defendance	+.27	Order	−.32
	Inferiorityavoidance	+.20	Understanding	−.27
			Affiliation	−.26
	Harmavoidance	+.19	Variability	−.19
Cluster 2	Defendance	+.27	Affiliation	−.29
	Harmavoidance	+.25	Nurturance	−.26
			Achievement	−.23
Cluster 3	Autonomy	+.33	Nurturance	−.34
	Sex	+.30	Understanding	−.32
	Exhibition	+.25		
Cluster 4	Dominance	+.30	Harmavoidance	−.30
	Counteraction	+.23	Abasement	−.25
Cluster 5	Sex	+.25	Nurturance	−.27
	Play	+.23	Order	−.25
Cluster 6	Aggression	+.34	Deference	−.34
	Rejection	+.26	Affiliation	−.27
Cluster 7	Sex	+.32	Abasement	−.29
			Succorance	−.28

tudes which combine with Order and Understanding (Affiliation, Nurturance and Deference) concern interpersonal relations in which Order and Understanding may be more difficult to achieve. The table makes possible other interpretations, some of which might be subjected profitably to experimental test.

SUMMARY

Two forms of the Picture Identification Test were given to two groups of 50 normal men and to two groups of 50 paranoid schizophrenic men. There were small but significant differences between the ages and the education levels of the combined normal and combined paranoid schizophrenic groups.

Two significant differences were found between the means of the single need-attitudes of normals and paranoid schizophrenics when Form I of the PIT was used. When Form II was given to two other groups, two other need-attitude means differed significantly. No single need-attitude was found that differentiated normals from paranoid schizophrenics on both test forms.

Correlations were found between all the possible pair-combinations of need-attitudes for each of the four groups taken separately. A critical ratio between the average correlations of normals and of paranoid schizophrenics did show statistically significant differences between the two groups. Need-attitudes which occurred most frequently in these combinations were, in rough order of frequency, Order, Achievement, Exhibition, Play, Affiliation, Aggression, Blamavoidance, Deference, Inferiorityavoidance, Rejection and Understanding.

Cluster analyses of the need-attitude score intercorrelations for normals and paranoid schizophrenics were compared. The clusters found showed some similarities between the normals and the paranoid schizophrenics, but they also showed sufficient differences to make possible tentative interpretations of the results.

REFERENCES

1. HOVLAND, C. I. Reconciling results derived from experimental and survey studies of attitude change. *Amer. Psychologist*, 1959, *14*, 8-17.
2. MURRAY, H. A. *Explorations In Personality*. New York: Oxford University Press, 1953.
3. PEAK, H. Attitude and motivation. In Jones, M. R. *Nebraska Symposium on Motivation*. Lincoln: University of Nebraska Press, 1955.

PSYCHOSEXUAL DEVELOPMENT IN TERMS OF OBJECT AND ROLE PREFERENCES

VLADIMIR PISHKIN

Research Laboratory, Veterans Administration Hospital
Tomah, Wisconsin

INTRODUCTION

An important problem in psychosexual development is the distinction between sexual *object* and sexual *role* preferences arising constantly in interpretation of projective findings, as well as in understanding of sexual factors in personality development and psychotherapy. Numerous concepts have attempted to differentiate between major characteristics of psychosexual development. Colley's[2] "sexual identity", "sexual biomode", and "sexual sociomode", Brown's[1] "sex-role identificaion", and Lynn's[6] "sex-role adoption" and "sex-role identification" are indicative of difficulties encountered in theoretical formulations within this area. A number of these concepts are ambiguous, not precise, and with overlapping meanings. The common factors underlying a majority of these propositions are based on the amalgamation of three major approaches to psychosexual development. The first orientation stresses Freudian[4] assertions on Oedipal situation and innate infantile sexuality. Mead[7, 8] and Horney[5] are primarily responsible for the second approach which places more emphasis on sociocultural factors. The third, more recent position emphasizes S-R conditioning of psychosexual behavior patterns[3, 9, 10].

It is proposed that psychosexual development is consistent with all these views. Generally, it is accepted that the nature of psychosexual development is influenced by parent-child interactions. There is evidence that our culture provides a relatively greater number of reinforcing situations for masculine behavior[6]. Our major assumption is that the *object* preference is determined by gratification of sexual needs as represented by Freud's concept of "object choice", while the development of *role* preference is influenced by masculine-oriented culture and sociocultural expectations acquired through S-R connections. It is plausible that both of these processes can be differentiated by an actuarial approach based upon conscious, self perceptions. The following hypotheses are made in relation to development of *object* and *role* preferences between the ages of 3 and 19:

1. *Ages 3-5.* Males show father *role* preference and females' *role* preference is for mother due to identification. Both males and females have *object* preference for mother through gratification of basic needs.

2. *Ages 6-12.* Through socialization, males begin to show less *object* preference for mother, i.e., mother as an *object* choice becomes less desirable, and the social expectations produce greater *role* preference for father. In females, the *object* preference shifts toward father and the *role* preference continues to be for mother, even though it is weakened by socialization.

3. *Ages 13-19.* Males' *object* preference for the opposite sex is externalized from the family, producing still lesser degree of *object* preference for mother with continuous *role* preference for paternal stereotype. Females weaken their *object* choice of father; their *role* preference for mother is lessened due to the effects of masculine-oriented culture.

METHOD

Subjects. *Ss* were 584 children between the ages of 3 and 19, obtained from nursery schools, grade schools and high schools of a small northwestern community. There were 292 males and 292 females divided into three age groups: Ages 3 to 5 (N = 144), 6 to 12 (N = 208), and 13 to 19 (N = 232). There was an equal number of males and females in each age group. The *Ss* within each age and sex groups were randomly divided into *object* (N = 292) and *role* (N = 292) subgroups. All *Ss* were selected

randomly with a restriction that they were living with both of their parents at the time of the investigation.

Procedure. Each *S* in the *object* subgroup responded to the question: "Whom do you like better, your mother (father) or your father (mother)?" Each *S* in the *role* subgroup was asked the following question: "Would you like to be like your mother (father), or like your father (mother)?" The first appearance of "mother" and "father" in the questions was equalized. Each *S* age 3 to 8 was interviewed individually. Their responses to the questions were recorded verbatim by the investigator. The *S*s, age 9 or older, were told that they were to write an answer to the question presented on a slip of paper. All the *S*s were told not to reveal their identity and only to indicate their age and sex. All responses were classified by preference for mother (*M*), no preference (*N*), i.e., preference for both parents or no response, and father preference (*F*) for each age, sex, *object* and *role* subgroups.

Results and Discussion

Chi square analysis was used to determine the significance of differences between *object* and *role* preferences within each age and sex groups. Cross-comparisons of sex, age, *object* and *role* subgroups were also made. Proportion of *object* and *role* preferences for each of the three age groups and for males and females are represented in Figs. 1 and 2, respectively.

Fig. 1. Distribution of Male Object and Role Preferences for Mother (*M*), No Preference (*N*), and Father (*F*) Categories.

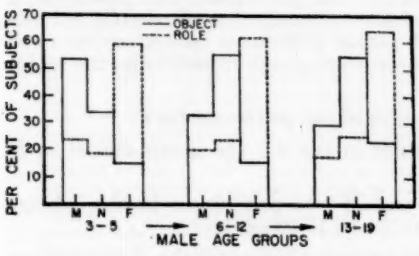
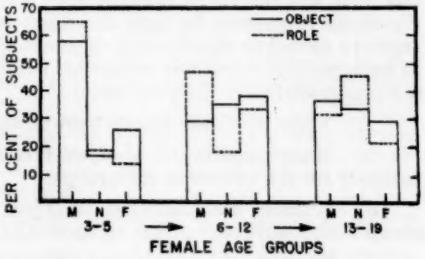


Fig. 2. Distribution of Female Object and Role Preferences for Mother (*M*), No Preference (*N*), and Father (*F*) Categories.



Males. *Object* and *role* preferences differentiated males of all age groups beyond the .001 level of confidence with the *role* preference of *F* category throughout all the age groups. In the 3-5 age group, the *object* preference was predominately for *M* ($p < .001$), while the *role* preference was for *F* ($p < .001$). The 6-12 and 13-19 age groups showed significant *N* choice in *object* preference with a predominate *role* preference for *F* ($p < .001$).

Females. The 3-5 age group demonstrated a consistent *object* and *role* preference for *M* ($p < .001$). In the 6-12 age group, however, the *object* preference shifted somewhat toward *N* and *F* preferences while the *role* preferences remained for *M* with some increase for *F* category. However, these shifts were not statistically significant. In the 13-19 age group, there was a trend of preference toward *N* category in both *object* and *role*. Again this differential was not significant, although a shift away from maternal stereotypy is suggested.

Within Groups. The *role* preference in males showed no significant shift over the three age groups, while the females demonstrated a significant change toward *N* category with increasing age ($p < .01$). The *object* preference of both males and females showed no significant trend over the age groups. Significant *role* preference differential was found between males and females of the 3-5 and 6-12 age groups with males' *role* preference of *F* and females' of *M* category ($p < .001$).

It is implied that socio-economic reinforcement and emphasis on material gain provided by our culture is such that it strengthens the desirability of masculine *role* (*F* preference). The results generally support the hypothesis that females' preference of *role* progresses from same sex to an opposite sex oriented world. However, inconsistent with the views of Lynn [6], it is apparent that differential *object* and *role* preferences of the sexes are in force as early as in the 3-5 age group.

The present findings may have wider implications in differential assessment of projection and identification. Sexual identification, as measured by human figure drawings, doll play, preferences for objects or pictures of objects, and other projective techniques, does not seem to provide a distinction between a need for satisfaction of biologically determined sexual needs (*object*) and learned preferences of behavior (*role*). It is noteworthy that information made available through conscious, self-appraisal can aid in discrimination of the two types of phenomena. Since projection is ascription of unconscious feelings unacceptable to one's ego, it has many elements in common with the *object* preference and Oedipal feelings. Identification is dynamically similar to the present concept of *role* preference, wherein certain traits and behaviors are desired or incorporated even though they are characteristic of the opposite sex. The general orientation in diagnosis and psychotherapy of psychosexual dynamics should be determined with the awareness of these two distinct aspects of personality which are *consciously* differentiated.

Summary and Conclusions

The study was designed to determine psychosexual *object* and *role* preferences of 584 male and female *Ss* between ages 3 and 19. The preferences were indicated by *S*'s choice of parents for both the *object* and the *role* categories. Conscious self-perceptions served to significantly differentiate parent preferences between sexes as well as between *object* and *role* categories within each sex group. Specifically, the major findings revealed by the data were:

1. Males of all age groups showed significant *role* preference for *F*.

2. Males' significant *M-object* preference in the 3-5 age group shifted to *N* category for the two older age groups.

3. Females demonstrated *M*-object and *M*-role preferences in the 3-5 age group with a shift to *N* and *F* categories at 6-12 and 13-19 age levels.

4. There were no significant differences between the sexes in the *object* category at the 3-5 age level (*M* preference); significant differences were found in the older two age groups—males showing preference for *N* and females being evenly distributed.

References

1. Brown, D. G. Sex-role preference in young children. *Psychol. Monogr.*, 1956, *70*, No. 14 (Whole No. 421).
2. Colley, T. The nature and origins of psychological sexual identity. *Psychol. Rev.*, 1959, *66*, 165-177.
3. Dollard, J., and Miller, N. E. *Personality and psychotherapy.* New York: McGraw-Hill, 1950.
4. Freud, S. *The basic writings of Sigmund Freud.* New York: Modern Library, 1938.
5. Horney, Karen. *The neurotic personality of our time.* New York: Norton, 1937.
6. Lynn, D. B. A note on sex differences in the development of masculine and feminine identification. *Psychol. Rev.*, 1959, *66*, 126-135.
7. Mead, Margaret. *Sex and temperament in three primitive societies.* New York: Morrow, 1935.
8. Mead, Margaret. *Male and female.* New York: Morrow, 1949.
9. Miller, N. E. Theory and experiment relating psychoanalytic displacement to stimulus-response generalization. *J. abnorm. soc. Psychol.*, 1948, *43*, 155-178.
10. Mowrer, O. H. *Psychotherapy: Theory and research.* New York: Ronald Press, 1953.

THE PSYCHOTIC REACTION PROFILE*

MAURICE LORR, JAMES P. O'CONNOR AND JOHN W. STAFFORD

The Catholic University of America

PROBLEM

The aims of this study were two-fold. The first goal was to construct, field test and standardize a multidimensional behavior inventory for the description of observable psychotic patient behavior in a hospital setting. The second goal was to test a set of hypotheses concerning psychotic reaction patterns that have been repeatedly identified in multiple factor studies of interview and ward behavior. If the more significant reaction patterns would be evidenced on the ward as well as in the interview situation, then an inventory could be developed for use by nurses and aides who actually have a greater opportunity for observing patient behavior than do the relatively few psychologists or psychiatrists on most hospital staffs.

The hypothesized primary constructs which the behavior inventory was designed to measure were Resistiveness and Motor Disturbances (catatonic-like behavior), Withdrawal vs. Sociability, Perceptual Distortions, Conceptual Disorganization, Melancholy Agitation, Paranoid Projection, Hostile Belligerence and Lack of Interest in Self or Others. In addition an interpersonal parameter of Dominance *vs.* Submissiveness was postulated. All patterns except the last have been isolated in at least two factor studies of rated interview and ward behavior. Four higher level or second order constructs were also hypothecated: Withdrawal, Belligerence, Thinking Disorganization and Intropunitiveness. (2, 5, 7)

METHOD

Construction of the Inventory. Published inventories, rating scales, reports of factor studies, and several psychiatric texts were scrutinized as sources for constructing some 400 behavior statements. A group of clinical psychologists familiar with psychotic patient behavior then allocated each statement to one of the nine defined areas of behavior disturbance previously described. Equivocal, vague, general, and duplicate statements were eliminated by this process. The final inventory consisted of 172 items distributed among nine subscales. The observers were instructed to record their observations of each patient on a separate answer sheet by indicating whether a statement was true, not true, or did not apply.

Collection of Data. Letters were sent to 150 psychiatric hospitals known to have psychologists or psychiatrists able and willing to collect the data in a responsible fashion. Each letter was accompanied by a detailed research outline, a set of instructions for collecting data and a specimen set of the Multidimensional Check List. The cooperating hospitals were asked to secure observations by their more intelligent and conscientious nurses and aides on a representative sample of patients. The only restrictions were that patients included be less than 55 years of age, ambulatory, and primarily functional psychotics. Observers selected were to be informed of the purpose of the project and to be given ample opportunity to familiarize themselves with the Check List and its instructions. Check Lists were to be completed under the supervision of a psychologist or psychiatrist at the close of a one week observation period.

Returns of 10 to 50 cases were received from a country-wide group of 47 hospitals in the United States and Canada. Of the 1,397 records received, 288 were dropped because of an excessive number of "Doesn't Apply" ratings; and others were found unusable for other reasons. Of the 958 cases remaining, 688 records were for male and 270 for female patients. The items were next examined. Some 31 items were eliminated because the per cent of "Doesn't Apply" ratings suggested that the behavior was too difficult to observe or simply was not available for observation. Seventeen additional items were dropped because of poor discrimination (variance less then .08), leaving 124 items for analysis.

Method of Scale Construction. Clinically it seemed reasonable to assume that the more severe the psychopathological reaction, the more probable that the deviant behavior described by a statement will be manifested. Thurstone has called such items "increasing probability" items. It follows that for a scale constructed from such items to measure a single underlying variable, if the score of patient A is higher than the score for patient B, then A is more severely disturbed than B. The method of homogeneous keying (4) is designed to select items of the increasing probability type to form a scale. The method can be used to evolve scales that are homogeneous, maximally

*This study was supported by research grant (M-1308) from the National Institute of Mental Health, Public Health Service.

FIG. 1. PSYCHOTIC REACTION PROFILE (PRP). SUBJECTS ARE ASKED TO RATE STATEMENTS AS TRUE, NOT TRUE OR DOESN'T APPLY. LETTERS IN PARENTHESES INDICATE SCALES.

1. Usually stays by himself. (W)
2. Tells the other patients what to do. (B)
3. Shows real sadness. (D)
4. Occasionally talks to himself. (TD)
5. Ignores the activities around him. (W)
6. Sometimes does the opposite of what he is asked to do. (B)
7. Seems to be unhappy. (D)
8. Answers sensibly when talked to. (TD)
9. Never says more than three or four words at a time. (W)
10. Acts as though the ward attendants are against him. (B)
11. Usually looks tired and all worn out. (D)
12. Sees and hears things that are not there. (TD)
13. Seems concerned about what others think of him. (W)
14. Resists suggestions and requests from aides. (B)
15. Seems scared all the time. (D)
16. Sometimes uses words that aren't understandable. (TD)
17. Doesn't take part in back and forth conversation. (W)
18. Resists treatment from the doctors. (B)
19. Usually looks worried and nervous. (D)
20. Spends a lot of time talking to himself. (TD)
21. Has to be pushed to follow routine. (W)
22. Acts as though the hospital is persecuting him. (B)
23. Talk is mostly not sensible. (TD)
24. Nearly always chatting with someone. (W)
25. Is likely to hit someone for no apparent reason. (B)
26. Usually knows what time it is. (TD)
27. Asks for things; doesn't wait for things to be given to him. (W)
28. Blames the hospital for lack of attention and care. (B)
29. Talks whether anyone is listening or not. (TD)
30. Is slow thinking and a little confused. (W)
31. Often shouts and yells. (B)
32. Talks to himself about imaginary or real faults. (TD)
33. Interested in nothing. (W)
34. Bosses the other patients. (B)
35. Sometimes giggles in a silly way. (TD)
36. Is able to talk about his own problems. (W)
37. Acts superior to other patients. (B)
38. Does not know the names of ward aides. (TD)
39. Is always doing something. (W)
40. Quick to fly off the handle. (B)
41. Repeats words and phrases in a meaningless way. (TD)
42. Shows no response to entertainment. (W)
43. Yells at attendant when he is dissatisfied. (B)
44. Makes faces and strange movements that do not make sense. (TD)
45. Has to be helped along to stick to any activity. (W)
46. Becomes easily upset if something doesn't suit him. (B)
47. Does not know where he is. (TD)
48. Says thanks when something is done for him. (W)
49. Becomes noisy and hilarious at times. (B)
50. It is difficult to understand what he is saying most of the time. (TD)
51. Never volunteers information about himself. (W)
52. Upsets other patients by the way he talks to them. (B)
53. Smiles a lot to himself without any sensible reason. (TD)
54. Acts dead to the world; doesn't seem to care what is going on. (W)
55. Has a sarcastic way of talking to other patients. (B)
56. Drifts off the subject when he talks. (TD)
57. Starts conversations with aides to become better acquainted. (W)
58. Often swears and uses obsene language. (B)
59. Often messy in eating habits. (TD)
60. Never asks for anything; waits for things to be given to him. (W)
61. Demands the attention of doctors. (B)
62. Usually is slow moving and sluggish. (W)
63. Often irritable, grouchy, or complaining. (B)
64. Is backward about talking to you. (W)
65. Doesn't swear or curse in the presence of doctors and aides. (B)
66. Speaks softly; often difficult to hear. (W)
67. Complains about the food and care he receives. (B)
68. Reads newspaper. (W)
69. Loses temper when dealing with other patients. (B)
70. Will do anything for recreation that comes up. (W)
71. Sometimes threatens to assault others. (B)
72. Doesn't mix with other patients. (W)
73. Has little interest in the problems of others. (W)
74. Asks for help from other patients when he needs it. (W)
75. Has no friend on the ward. (W)
76. Tries to be friendly with other patients. (W)
77. Laughs if he is kidded. (W)
78. Would sit all day unless directed into activity. (W)
79. Is good company. (W)
80. Laughs or smiles at funny comments or events. (W)
81. Likes to go for exercise. (W)
82. Shows real friendliness towards at least one other patient. (W)
83. Seems always busy with plans and projects. (W)
84. Shows occasional interest in news and current events. (W)
85. Seldom listens to radio or watches TV. (W)

differentiating and minimally intercorrelated. Since the method seemed suited for the task at hand it was utilized in parallel analyses of item covariance matrices for a sample consisting of 500 male records and 250 female records.

In the initial trials the nuclei chosen to evolve independent subscales consisted of three items selected from each of the nine behavior disturbance areas postulated. However, it was soon apparent that the subscales evolved were too highly correlated although they centered on the presumed primary factors. The next set of trials were centered on four or five variables representative of the four higher level constructs previously described. Four essentially independent scales comprising the Psychotic Reaction Profile given in Fig. 1 were successfully generated by this last procedure—Withdrawal, Thinking Disorganization, Paranoid Belligerence, and Agitated Depression.

Results

Interpretation of the Scales. The Withdrawal Scale, which consists of 38 statements, is the best balanced and the most reliable of the scales developed. The scale score distribution is relatively normal as contrasted with the strongly positively skewed distributions of the other three scales. It appears to be very similar to the second order Withdrawal factor identified in at least three factor studies. The Withdrawal Scale appears to be bipolar and is defined by three discernible elements: activity level, manifest interest in the environment, and interpersonal responsiveness. One pole of the scale defines a withdrawal reaction consisting of a lack of interest or apathy, a failure to respond to others, and a reduced activity level. The opposite pole is descriptive of an excessive busyness, overactivity, and a high level of interpersonal reactivity.

The Thinking Disorganization Scale (18 items) defines a broad parameter probably central to schizophrenia. Disorientation, irrelevant and incoherent speech, hallucinations, and peculiar movements are characteristic of this pattern. The scale appears to measure a reaction probably identical with a similarly labelled second order factor isolated in several independent studies[5, 7, 8].

The third scale (24 statements) defines a Paranoid Belligerence. Patients scoring high on this scale are hostile, irritable, resistive, bossy, and paranoid. To some clinicians this pattern would represent a hostile acting out. A very similar cluster called Belligerence was first isolated by Lorr, Jenkins, and Holsopple[6] and again in Wittenborn's data[5]. Recently Bostian, Smith, Lasky, Hover, and Ging[1] reported a very similar parameter in an analysis of a mental status examination.

The fourth and final scale consists of only five items. The anxiety and obvious depressive elements that characterize the scale suggest an Agitated Depression.

Table 1. Intercorrelations, Reliabilities and Mean Scores for Open and Closed Ward Patients on Form Scales

Scale Correlations*	Thinking Disorg.	Withdrawal	Par. Belligerence	Agitated Depression
Thinking Dis.	(.90)	.38	.26	.06
Withdrawal	.40	(.94)	.10	.36
Par. Belligerence	.37	.15	(.90)	.17
Agitated Depress.	.05	.30	.12	(.74)
Inter-rater reliability				
Male patients	.88	.90	.85	.65
Female patients	.81	.92	.78	.58
Sample A means				
Open ward	2.74	13.62	4.31	1.36
Closed ward	4.26	17.22	3.30	1.74
Sample B means				
Open ward	2.58	15.09	3.96	1.44
Closed ward	4.06	17.03	4.79	1.82

*Correlations for male patients above diagonal; correlations for female patiente below diagonal

Because of its brevity this scale is the least reliable and discriminating. It is judged to be identical with the Agitated Depression isolated by Wittenborn[9], Lorr, *et al.*,[7], and Guertin[3].

Table 1 presents the intercorrelations between the scales. The two highest correlations are between Thinking Disorganization and Withdrawal, and between Withdrawal and Agitated Depression. As may be seen all of the scales have a satisfactory degree of independence.

Reliability. The Kuder-Richardson formula #20 estimates of scale reliability are shown in Table 1. All coefficients except that for Agitated Depression are at least .90. To ascertain inter-rater reliabilities, data were obtained from a state hospital on a group of 54 male and a group of 52 female psychotics each rated by two attendants. The correlations shown in Table 1 indicate that all scales except the agitated depressive scale have a satisfactory reliability. In fact such reliabilities compare favorably with those reported for well constructed achievement and aptitude tests.

Construct Validity. Several sources of evidence for construct validity are available. A proposed measure might be expected to have a content consistent with the proposed interpretation. The preceding description of the scales supports the hypotheses suggested from the standpoint of content. The correlations of the proposed measures with external criteria should not all be zero; some should be substantial and consistent with expectations. For example, the scales should differentiate open from closed ward patients insofar as open ward patients are likely to be less severely ill than patients locked on their wards. To ascertain the evidence for this criterion, two independent samples of 55 male open and 110 male closed ward patients were compared on each of the four subscales. Table 1 presents the mean scores obtained for the open and closed ward status patients for the two samples. Thinking Disorganization is significantly higher (one-tail test) for closed ward patients in both samples at a probability less than .025. Withdrawal is significantly higher for the closed ward patients at the 10 percent level in Sample A and at the .025 level in Sample B. Agitated Depression separates the two groups in expected directions at the 10 percent level in both samples. Paranoid Belligerence, on the other hand, fails to differentiate the two groups in either sample. One plausible explanation is that patients who score high on the subscale are stabilized paranoids, fairly well integrated who are typically allowed considerable freedom on hospital grounds. The above findings probably under-represent the differentiating power of the scales. No definition of "open ward" was provided the hospitals. In addition it seems highly likely that there were marked differences between hospitals in policies concerning the degree of freedom permitted patients.

To test whether unique patterns on the four scales could differentiate open from closed ward patients, each scale was dichotomized at or near the third quartile. Since the four scales were dichotomized, 16 unique patterns resulted. Each patient was allocated to one and only one pattern. A test was then made of the utility of the patterns on a sample of 110 open ward patients and 220 closed ward patients. Of the closed ward patients 80 percent were correctly identified. However, the open ward patients could not be identified any better than chance. These findings suggest that with further research the scale patterns would have considerable utility for ward assignment.

Sex and Age Differences. The initial keying procedure, applied separately to the covariance matrices based on the descriptions of male and female patients, revealed no differentiating items. While the order of inclusion of items in a cluster was not always the same, the final keys were essentially identical for both sexes. The norm data tend also to support the belief that the psychotic reaction patterns do not differ greatly in men and women. Each of the four scale score distributions was examined for sex differences. The only difference obtained, significant at the .001 level or better, was on Paranoid Belligerence. Women patients are rated as more belligerent and hostile than male patients. Possibly this form of emotional expression is more acceptable to

ward personnel and less likely to be suppressed or penalized when exhibited by women than by men.

The scale scores for the male patients were also examined for age differences. Correlation ratios were computed between age and each of the subscales. The age range included in the analysis was from 20 through 52 with the modal value at age 36 for each of the distributions. Only the correlation (.15) of age and Thinking Disorganization was significant at the five percent level. The thinking disorganization of patients, as measured by the scale, increases with age. For the other scales the median age remains invariant over the various score intervals.

Summary

Descriptions of 1400 psychotic patients in 47 hospitals were recorded on a 172 item behavior inventory called the Psychotic Reaction Profile (PRP). The method of homogenenous keying was applied to evolve four scales of 85 items. The scales labelled Withdrawal, Thinking Disorganization, Paranoid Belligerence, and Agitated Depression, have a satisfactory rater reliability and internal consistency. All except the Belligerence scale discriminate between open and closed ward patients. Centile norms by sex and age are available for each scale. In addition, norms are available for 16 unique patterns.[1]

[1]From Dr. Maurice Lorr, Psychology Department, Catholic University, Washington 17, D. C.

References

1. Bostian, D. W., Smith, P. A., Lasky, J. J., Hover, G. L. and Ging, R. J. Empirical observations on mental status examination. *Arch. gen. Psychiat.*, 1959, *1*, 253-262.
2. Degan, J. W. Dimensions of functional psychosis. *Psychometr. Monogr.*, 1952, No. 6.
3. Guertin, W. A factor analytic study of schizophrenic symptoms. *J. consult. Psychol.*, 1952, *16*, 308-312.
4. Loevinger, Jane, Gleser, Goldine C. and DuBois, P. H. Maximizing the discriminating power of a multiple-score test. *Psychometrika*, 1953, *18*, 309-317.
5. Lorr, M. The Wittenborn psychiatric syndromes: An oblique rotation. *J. consult. Psychol.*, 1957, *21*, 439-444.
6. Lorr, M., Jenkins, R. L. and Holsopple, J. Q. Factors descriptive of chronic schizophrenics selected for the operation of prefrontal lobotomy. *J. consult. Psychol.*, 1954, *18*, 293-296.
7. Lorr, M., Jenkins, R. L. and O'Connor, J. P. Factors descriptive of psychopathology and behavior of hospitalized patients. *J. abnorm. soc. Psychol.*, 1955, *50*, 78-86.
8. Lorr, M., O'Connor, J. P. and Stafford, J. W. Confirmation of nine psychotic symptom patterns. *J. clin. Psychol.*, 1957, *13*, 252-257.
9. Wittenborn, J. R. Symptom patterns in a group of mental hospital patients. *J. consult. Psychol.*, 1951, *15*, 290-302.

A WARD BEHAVIOR RATING SCALE FOR MENTAL HOSPITAL PATIENTS[1]

E. I. BURDOCK AND G. HAKEREM
ANNE S. HARDESTY JOSEPH ZUBIN

Biometrics Research, New York State Department of Mental Hygiene and Department of Psychiatry, Columbia University

Introduction

A behavioral rating of 150 items has been designed to measure severity of illness and an index of response to treatment for mental patients, both new admissions and chronic residents.

Material was selected by reference to the literature of psychopathology and by consultation with representatives of the various mental health disciplines. The present effort capitalizes on previous work such as The Behavior Chart[2] inspired by Adolph Meyer and still in use at the Phipps Psychiatric Clinic; the Rating Scheme for Conduct[6]; A Schema for the Quantitative Measurement of Abnormal Emotional Conditions[5]; the Wittenborn Psychiatric Rating Scales[7]; Lorr's[3] Multidimensional Scale for Rating Psychiatric Patients; and the Hospital Adjustment Scale of Ferguson, McReynolds, and Ballachey[4].

The items[2] were constructed to reflect observable units of behavior such as facial expression and grooming; eating activity and toilet habits; physical status; attitudes; cooperativeness; communication; vocalization and speech patterns; interpersonal relations; hostility or aggressiveness; mannerisms; affect; and special symptoms. No attention is directed at rating a patient's underlying dynamics, the whole effort being concentrated on determining current status.

Ambiguity was minimized and intelligibility increased by choice of the simplest and most direct wording for each item. Technical terms and explanatory examples, where necessary, were segregated by parentheses. All items were tested for clarity and specificity by try-out with nurses and attendants. The items of behavior have been drawn from three categories of observation: appearance and deportment, behavior in verbal contexts, and adaptation to ward routine.

Procedure

The items are marked by circling "T" (for True when the patient displays the indicated behavior) or "NT" (for Not True when the patient does not behave in the manner described). To maximize the amount of information, a forced choice technique has been employed, *i.e.*, the rater has been directed to answer every item as "True" or "Not True" to the best of his ability. Items concerned with adaptation to ward routine may require reference to the nursing record if the opportunity for observation has not arisen. The rating has been focussed on the patient's *behavior during the preceding week*. Historical material or facts known about the patient prior to this time are excluded from consideration.

Prospective raters should have a training session before being left on their own. After a little experience, about ten or fifteen minutes has generally been enough time to complete the whole scale for a patient. A dotted line has been placed after each item to be used for any qualifications which the rater has wanted to make in addition to rating the item as "T" (True) or "NT" (Not True). Items have been scored zero or one according as the behavior is adaptive or not. Maximum score is thus 150.

---

[1]This research has been supported by Research Grant M-3546 from the National Institute of Mental Health of the National Institutes of Health, Public Health Service.

[2]Interested readers may secure copies of the Ward Behavior Rating Scale for research purposes by writing to the authors.

Reliability

Two ratings were made for each patient by regular ward nurses or attendants after a training session to impart instructions and stimulate motivation. Normality was tested by fitting to chi square the ratios of the variances within subjects to the mean variance within subjects.

Because of work schedules different observers were paired across patients and the same observer appeared sometimes as first, sometimes as second rater. Therefore, inter-judge agreement was analyzed by intra-class correlation. An intra-class correlation of .67 was obtained from a one-way analysis of variance applied to paired ratings on 107 patients at New York Psychiatric Institute.

A similar analysis of a sample of 50 pairs of ratings from a New York state mental hospital yielded an intra-class correlation of .84. The correlation dropped to .40 for 61 pairs from a VA hospital. On a sample of ten patients from another New York state hospital the same four judges rated each patient allowing a two-way analysis of variance which provided an intra-class correlation of .77.

Internal consistency was studied by applying Hoyt's model for a two-way analysis of variance to a composite sample of 177 protocols from several state hospitals. A significant F between items supported the hypothesis that the items were making independent contributions to the total score. The coefficient of internal consistency was .95; the standard error of an individual score was 4.4.

Validity

Criterion studies with preliminary pools of items contributed two items which identified young males released before the end of three months and 19 items which distinguished young males out of the hospital at the end of 12 months.

A preliminary form of the scale was used in conjunction with a study[1] of the effectiveness of a special geriatric unit at Central Islip State Hospital. The rating scale proved effective both by showing that the experimental and control groups were comparable at intake and by reflecting a significant ($p < .01$) improvement in the experimental group on four-week follow-up.

A study, with the present form, on female geriatric patients, distinguished a group treated with drugs from a matched placebo group at the .05 level of significance. The predictive validity of the rating scale has been tested against an Outcome Index which combines length of hospitalization and number of admissions and releases into a single composite criterion ranging from zero to one. The Outcome Index computed for a six-months follow-up on 107 patients at New York Psychiatric Institute showed a low but significant correlation ($r = .21$; $p < .05$) with scores on the Ward Behavior Rating Scale.

Summary

A behavioral rating scale of 150 dichotomous items has been designed to provide a measure of severity of illness and an index of response to treatment for mental patients.

References

1. Burdock, E. I., Elliott, Helen E., Hardesty, Anne S., O'Neill, F. J., and Sklar, J. Biometric evaluation of an intensive treatment program in a state mental hospital. *J. nerv. ment Dis.*, in press.
2. Kempf, E. J. The Behavior Chart in mental disease. *Amer. J. Insanity*, 1915, *71*, 761-772.
3. Lorr, M. Rating scales and check lists for the evaluation of psychopathology. *Psychol. Bull.*, 1954, *51*, 119-127.
4. McReynolds, P., Ballachey, E., and Ferguson, J. Development and evaluation of a behavioral scale for appraising the adjustment of hospitalized patients. *Amer. Psychologist.*, 1952, *7*, 340.
5. Moore, T. V. The Essential Psychoses and their fundamental syndromes. *Studies in Psychology and Psychiatry* III. Baltimore: Williams & Wilkins, 1933.
6. Plant, J. S. Rating scheme for conduct. *Amer. J. Psychiat.*, 1922, *1*, 547-572.
7. Wittenborn, J. R. A new procedure for evaluating mental hospital patients. *J. consult. Psychol.*, 1950, *14*, 500-501.

RELIABILITY-UNRELIABILITY OF ANCILLARY PSYCHIATRIC EVALUATIONS[1]

MILTON B. JENSEN WILLIAM E. MORRIS

Veterans Administration Hospital, Salisbury, North Carolina

Introduction

Research and treatment at this NP hospital often resort to psychiatric evaluations by others than psychiatrists and psychologists. Mostly they are made by nurses and nursing assistants. Sometimes the evaluations employ such devices as Lorr's Multidimensional Scale[6], or such behaviors as frequency of vocalization, etc. Systematic analyses, conducted by us over a period of several years, have revealed a wide range of agreement-disagreement between ancillary judges simultaneously evaluating the same patients on the same psychiatric continuum. Intercorrelations (Rhos and Pearson *rs*) between them have averaged in the neighborhood of .55. Some have been negative and a few have been above .90. (Intra-class correlations and Kendall coefficients of concordance for the same data, though generally slightly higher, were not significantly so. Examples: For one set of rankings—22 patients by 8 judges—average Rho was .64 and W was .67. 5-point scale ratings on the same patients by the same judges gave an average r of .52; r_{cc}, .55. We treated our data by the Spearman and Pearson methods and Shen's formula[7] because this gives more precise reliability coefficients for each of a group of evaluators than can be secured non-parametrically.) Our findings indicate that only a small percentage of nurses and nursing assistants render reliable psychiatric judgments of NP patients. We have found no easy method of separating these from the majority who do not do so. We doubt that any reasonable amount of training would make satisfactory ancillary psychiatric evaluators of some. But support for the idea that such ability *per se* is fairly consistent can be shown by the following finding[5]: Each of seven nurses, with equal opportunity for observation, evaluated all the patients on two wards. Rho between nurse reliability co-efficients on one ward as compared with those for the second ward was .82 ($P < .05$).

Radical disagreement between evaluators should be expected when they judge in areas in which they are not primarily trained. Hunt and Jones[1] have shown that undergraduates are able to deal with a global concept of "schizophrenicity" but are not as able as trained clinicians to judge separate dimensions of schizophrenic thought. Also, in this day of more favorably staffed hospitals with pharmacies filled with tranquilizers, it is a difficult task for anyone to assess patients reliably on such a continuum as "agitation", even if this concept were made more operational. Our findings document the need to select, train, and supervise evaluators and to objectify the continua we ask them to deal with. And they induce skepticism of researches, scales and treatments based on psychiatric evaluations by others than psychiatrists and psychologists unless the judgments of several are combined to give high reliability, and there has been cross-validation and replication. That many of our evaluators did poorly is primarily a reflection on the professional staff, including the present authors. That so many did so well is surprising indeed.

We have not found comparable disagreement between our psychiatrists and psychologists in making such judgments, nor between non-psychiatric evaluations by administrative personnel.[1] In one study, psychiatric evaluations between a psychologist and three psychiatrists correlated .80[5]. In another, two psychiatrists correlated .91 in rating degree of pathology of patients[3], and, in a third study using a non-psychiatric scale, the rankings of housekeeping maintenance men by their supervisors correlated, on the average, .79[4].

---

[1]Ancillary refers to psychiatric evaluations by others than psychiatrists and psychologists. We call them "psychiatric" when they are based on a psychiatric continuum such as depression or suspicion. Non-psychiatric evaluation refers typically to evaluation of observable behavior.

Procedures and Results

Interrelationships between the ancillary psychiatric evaluations obtained in two separate studies[3, 5] are tabled in this paper, together with reliability coefficients derived from the intercorrelations by Shen's formula[7]. In one study, both nurses and nursing assistants evaluated their patients; in another study, only the nursing assistants. It is significant that, in the studies cited in this paper, all the scales on which disagreement between ancillary judges was extreme asked for psychiatric judgments: "How agitated is the patient? Does he seem to be unhappy? Does he ignore what goes on around him?" Very few of the items on these scales called for simple, observable data such as how many times the patient spoke during a given period.

In interpreting the intercorrelations and reliability coefficients in Table 1 it should be remembered that the former can range from -1.00 to $+1.00$ but that reliability coefficients cannot be negative. The reliability coefficients may be used for comparisons between the accuracy of different judges irrespective of whether they assess the same number of, or even the same subjects[7], while the intercorrelations measure only agreement-disagreement. The data are more meaningful if one or two common relationships are kept in mind: The correlations between the heights of fathers and grown sons and between the school grades of parents and children are about the same as the .55 average of 466 intercorrelations from one of our studies[5]. While such correlations are well above chance (.00), taken singly, they have very limited predictive power and must be replicated or combined with other measures to

Table 1. Intercorrelations, Reliabilities and Mann-Whitney p-Values for Evaluations by Nurses and Nursing Assistants

Value	Intercorrelations* (N = 466)	Reliabilities* (N = 146)	P** (N = 49)
1.00	3	2	7
.95	5	1	1
.90	8	6	2
.85	14	1	6
.80	25	16	5
.75	27	10	7
.70	53	20	4
.65	45	11	7
.60	37	14	1
.55	62	10	3
.50	45	9	4
.45	32	11	2
.40	21	10	
.35	24	5	
.30	20	3	
.25	9	4	
.20	10	4	
.15	11	6	
.10	5		
.05	3	2	
.00	3		
−.05	3		
−.10			
−.15			
−.20			
−.25	1		
Average	.55	.57	.75*

*The intercorrelations and reliabilities are from a study of the evaluations of 130 patients by 14 nurses and 36 nursing assistants[5]. The *p*-values are from the ratings of 58 patients by 49 nursing assistants.

***p*-values give the chances (in portions of unity) that the raters were describing the same thing or things. They are not identical with correlation coefficients.

give reasonably accurate placement of individuals in rather gross categories such as school grades, major diagnostic categories, etc.

Table 1 shows that very few of our nurses and nursing assistants made psychiatric evaluations with reliability high enough for other than group differentiation. Only seven of 49 agreed at the .02 level of confidence as measured by the Mann-Whitney Test in one study [3] and only nine of 146 reliability coefficients from another study [5] were .88 or higher. Considering the .57 average of the 146 reliability coefficients (Table 1): A simple statistic (Spearman-Brown Formula) predicts that it would take the combined evaluations of seven such judges to produce reliability .90. Were we unfortunate enough to select evaluators no more reliable than the .175 average of the lowest 16, however, the combining of 60 such judgments would be needed to obtain a reliability of .90. One ("terminated") study at this hospital predicated its findings upon the psychiatric judgments of two nursing assistants who correlated −.21, and, another study of paired patients hinged on psychiatric evaluations by three nursing assistants whose judgments, on the average, intercorrelated only .47 (combined reliability, .73).

We found unstable relationship between the reliabilities of many ancillary evaluations made on the same patients at different times. In one study [5], Rho between the reliabilities of 15 replicated rankings (after six weeks) was .52. For 5-point scale reliabilities it was −.08. Furthermore, some of our ancillary judges could not agree with themselves when evaluating the same patients on the Agitation-Apathy continuum by two methods [6]. The correlations between their rankings and 5-point scale ratings ranged from .02 to .98 (Mean, .79).

Our data show that reliable ancillary psychiatric evaluations can be secured via psychiatric scales by using enough judges. But, without statistical analysis we would not know how many were enough and which ones should evaluate. The average reliabilities with which nurses and nursing assistants first ranked the patients on five wards on the Agitation-Apathy continuum [6] were .78, .70, .56, .48 and .46 (Range, .16— .90). Using the 5-point scale their average reliabilities were .52, .58, .71, .65 and .44 (Range, .30— .99). Fortunately, we secured evaluations from 8, 6, 7, 8 and 6 judges on these various wards. We estimate the reliabilities of combined judgments (all rankings and 5-point scale ratings) on these wards at .98, .97, .95, .94 and .91 (Spearman-Brown Formula). The number of separate judgments combined in these predictions are 16, 12, 14, 16 and 12, since each of the evaluators on each ward employed rankings and 5-point scale ratings.

The Problem of Validity

Even if ancillary psychiatric evaluations were adequately reliable, their validity would still be subject to question. While some such judges are astute at this task, most of them disagree radically with each other and with themselves. Obviously, with such disagreement not many of them can be right. Our data indicate that while the pooling of many such judgments results in something that is stable, we do not know what that something is. In one study, for instance, we found an average correlation of .72 between pathology ratings by two psychiatrists and by 49 nursing assistants. But, seven of the correlations were below .20 and one of them was −.44. And, the correlation between reliability and agreement with the psychiatrists was .07.

In all of our studies clinical psychologists and psychiatrists agreed well with each other and were reasonably reliable in evaluating via psychiatric scales. Intercorrelations between them have ranged from .64 to .94, with an average of .80. On the average, two of our psychology-psychiatry staff evaluate via psychiatric scales with combined reliability of .90. We have not found comparable agreement on psychiatric judgments by other disciplines. We recommend, therefore, that ancillary evaluations via psychiatric scales be validated (experimentally) against similar judgments by two psychiatrists, two psychologists or one psychiatrist and one psychologist.

Securing Reliable Ancillary Psychiatric Evaluations

In the light of our findings we might reasonably recommend that ancillary psychiatric evaluations be used only after competence of the judge has been experimentally established. If this were done, only a small percentage of nurses and nursing assistants would qualify to make such evaluations. This is to be expected. We see it as no more reasonable to expect high quality of psychiatric judgment without primary training in that area than to expect surgical skill to develop from association with surgeons and their patients. Perchance, as we see it, the training and experience of nurses and nursing assistants properly disqualifies them as psychiatric evaluators. To the extent that they are superior nurses and nursing assistants they are dedicated to other than evaluative work. Our data show, however, that ancillary behavioral ratings may have high fidelity. We suggest again that ancillary evaluations of NP patients be in terms of objective observable behaviors and that psychiatric interpretations of these data be made only by the psychiatric sophisticate. From our studies[2, 3, 4, 5] we have derived the following conclusions relative to the securing of ancillary evaluations for the use of psychiatrists and psychologists in researches and treatments.

1. Ancillary evaluators should be at least average in language mastery. In one of our studies[3] the 28 nursing assistants who made no clerical errors averaged higher, though not significantly so, on the Thorndike-Gallup Vocabulary Test than did the 18 who made the most errors (Means, 10.1 and 9.3, respectively), and the correlation between Thorndike-Gallup score and frequency of unwarranted use of the term "Doesn't Apply" was $-.34$ ($N = 47$). Also, Rank-Biserial correlation between frequency of unwarranted use of the term "Doesn't Apply" and the nursing assistant's agreement with the pathology ratings of two physicians on 32 patients was $-.33$. In our studies, very high vocabulary scores seemed not to further clerical accuracy, reliability or validity of nursing assistants' evaluations. But, a Thorndike-Gallup score below 9 (12 of 47 evaluators so scored) seemed almost disqualifying.

2. Ancillary evaluators should be considered *good* but not *very good* by their supervisors. In clerical accuracy, freedom from unwarranted use of the term "Doesn't Apply", reliability and agreement with physicians' ratings of severity of pathology, we found a definite trend for nursing assistants rated *good* by their nursing supervisors to excel those rated *very good* and for those rated *very good* to excel those rated *fair*[3]. Seemingly, the ability to impress supervisors, up to a certain point, is predictive of satisfactory performance as an evaluator. In the halo area it is probably a liability.

3. Ancillary judges can spend too much time in direct contact with (helping, supervising, etc.) the patients they evaluate. In one of our studies[3], the evaluations by nurses and nursing assistants were more satisfactory on the midnight to 8 a. m. shift than on the 4 p. m. to midnight shift and least satisfactory from 8 a. m. to 4 p. m. We think that prolonged, uninterrupted supervision of mental patients impairs ability to evaluate them by reason of at least the following factors: (a) boredom, (b) fatigue, and (c) loss of objectivity through the development of personal bias. The night shift personnel are less subject to these impairments. Theoretically, the lights go out at 9:15 p. m., and the patients are awakened at 6:15 a. m. Thus, the night raters observe only about two hours of "waking behavior". It should not be forgotten that significant behaviors occur at night. Certain patients become sleepless, agitated, fearful, hallucinated, aggressive, etc., and such significant and deviant behaviors can be readily observed because of seeing patients individually instead of in a group setting. (We incline to the opinion that more satisfactory ancillary evaluations could be secured from those specially assigned to this task.)

4. Ancillary evaluator reliability does not improve with practice. The factors leading to improvement of evaluations probably are identical with learning in any other area. Practice *per se* is not one of them. Our evaluators were no more reliable on any succeeding series of judgments than on the first groups they evaluated. And,

the correlation between reliability and number of patients evaluated (from 13 to 80) was only .02 for an N of 148 [5].

5. Length and breadth of experience beyond the orientation level did not insure ancillary evaluator reliability. In one of our studies [5] reliability correlated only .11 with number of months of continuous employment and .01 with number of treatment areas of the hospital in which the evaluator had worked.

Summary

This report summarizes six years' study of the reliability of scaled psychiatric evaluations. Mostly, it is concerned with judgments by nurses and nursing assistants of NP hospital patients. The statistical findings, and our non-quantified experience, lead us to the following: Adequate reliability of psychiatric evaluations cannot be assured without, (a) selection based on statistical analysis of like judgments by those from whom only a few evaluators are to be selected, or (b) combination of the judgments of several evaluators as one does in combining items and parts in standardizing psychological and educational tests. In general, we secured the same reliability from seven nurses and/or nursing assistants that we obtained from two of the psychiatry-psychology team. Even if ancillary psychiatric evaluators render reliable judgments, extreme care must be taken that they maintain the intended frame of reference, that they are motivated (and rewarded) for the task, and that they have sufficient time and freedom from other job pressures.

References

1. Hunt, W. A. and Jones, N. F. Clinical judgment of some aspects of schizophrenic thinking. *J. clin. Psychol.*, 1958, *14*, 235-239.
2. Jensen, M. B. Ward behaviors of 50 NP hospital patients related to viewing a typed movie and to time of day. V. A. Hospital, Salisbury, N. C., 1954. (Mimeographed)
3. Jensen, M. B., Morris, W. E. and Kandilakis, J. Inter-and intra-relationships of nursing assistants ratings of NP hospital patients. V. A. Hospital, Salisbury, N. C., 1957. (Mimeographed)
4. Jensen, M. B., Morris, W. E. and Mayberger, Z. Rated characteristics of hospital housekeeping maintenance men. V. A. Hospital, Salisbury, N. C., 1958. (Mimeographed)
5. Jensen, M. B., Sharp, W. T. and Higbee, D. S. The effects of stelazine on extremely agitated and extremely inert schizophrenic patients. V. A. Hospital, Salisbury, N. C., 1959. (Mimeographed)
6. Lorr, Maurice. Multidimensional Scale for Rating Psychiatric Patients. Veterans Administration, Washington, D. C., Technical Bulletin 10-507, November 16, 1953.
7. Shen, E. The reliability coefficient of personal ratings. *J. educ. Psychol.*, 1925, *16*, 232-236.

THE RELATION OF PERSON NAMING AND TENSION TO PROGNOSIS AND ADJUSTMENT OF CHRONIC MENTAL PATIENTS[1]

JEFFREY CAINE[2] AND GERDA RAND FILLENBAUM[3]

St. Louis State Hospital *St. Louis State Hospital and Washington University*

Problem

This study attempts to find a way of categorizing chronic mental patients that relates to prognostic outcomes and levels of hospital adjustment. Leads to prognostically relevant ways of categorizing such patients have been reported by other writers. Skottowe [4] and Thomas [5] found clinically that inability either to formulate "conceptual thoughts" upon personal topics or intelligibly to discriminate gradations of one's emotions was prognostically unfavorable. Peters [2] interpreted his results to indicate that a state of "integrity and dominance of the cortical centers" and a "generally heightened activity level" were prognostically favorable.

[1]The authors gratefully acknowledge the invaluable assistance of Bongky Lee and Helen Walker in the collection of data.

[2]Now at Worcester State Hospital, Worcester, Massachusetts, where his share of the investigation was completed as part of a USPHS postdoctoral fellowship in clinical psychology.

[3]Now at Chapel Hill, North Carolina.

Two somewhat similar ways of categorizing were investigated: (a) The ability of a patient to name persons was determined by previously observing her regular eating companions, and then finding if she could accurately name any of them when requested to do so in an interview. (b) The tension of a patient was determined during the interview by judging the height of the level attained by any marked and unexpected change of either body movement, vocal tone, or affective loading of language. The purpose of the study is to discover the relation of these indices of person naming and tension to prognostic outcomes and levels of hospital adjustment.

METHOD

Subjects. The *E* individually interviewed all the 71 female, white, mental patients on one state hospital ward; 63 of them were included in the study since *E* was able to classify them on person naming (referred to as naming) and rate them on tension. The other 8 *Ss* comprised 4 who were insufficiently cooperative, and 4 who ate alone and therefore could not be classified on naming.

The primary diagnosis was used for the classification of *Ss* with a mixed diagnosis. Where necessary the diagnostic terms were translated to conform to the current nosology of the American Psychiatric Association's *Diagnostic and Statistical Manual*[1]. Forty *Ss* had schizophrenia: 17 paranoid, 8 hebephrenic, 8 chronic undifferentiated, and 7 catatonic; 10 were lobotomized. Two *Ss* had paranoid states. Seven *Ss* had chronic brain syndrome (CBS) with convulsive disorder, 4 also had psychosis; 2 CBS with central nervous system syphilis; 1 CBS with cerebral arteriosclerosis; 1 CBS with psychosis; and 1 CBS with mental deficiency. Nine *Ss* had mental deficiency, 7 with psychosis.

Ages ranged from 20 to 71 with a mean of 49.7 and a median of 52 years.

Educational attainment was divided into six levels. Nine *Ss* terminated in grades 1 to 4 and/or attended a special school, 15 terminated in grades 5 to 7, 21 in grades 8 to 11, 8 in grade 12 including 1 who terminated in grade 9 and went to business college, 7 either had one year of university training or attended business college after reaching grade 11 or 12, and 3 had two or more years of university training—1 being a graduate.

Duration of institutionalization was calculated from the date of the first known institutionalization, disregarding any time spent outside an institution since then. Durations ranged from 1 to 37 years with a mean of 18.1 and a median of 18 years. Only one *S* had a duration of less than 3 years.

Durations of stay on the ward ranged from 2 weeks to 14 years with a mean of 4.7 and a median of 4 years; 15 had been on the ward for less than 1 year. This duration was calculated from the most recent date of entry to the ward, disregarding absence for general hospital services in the calculation.

Interview. All *Ss*, except for ten, were interviewed in random order in a room on their ward. The 10 *Ss* comprised 6 who were working off the ward, 2 who were seen at the end as they were inclined to chatter, and 2 who were in bed with illness. These *Ss* were seen after all the others.

The *E* introduced herself and an observer to *S* and said that they were from the psychology department. Then *E* said: "You may have seen me on the ward last week. Now I want to talk with each one of you so that we may be better able to help you all. Today I'd like to ask you a few questions. I want your own answers to them. *There is no right or wrong way of answering them.*" First, *S* was asked to guess how many patients were then living on the ward, and to name as many of them as possible (up to ten). Then, *S* was encouraged to name the attendants, aides and doctors who worked on the ward. Finally, *S* was asked to say with whom she ate or sat the previous week, with whom she would and would not like to eat or sit, and which patients she thought would and would not like to eat or sit with her.

As far as possible each *S* was asked all the questions. If an *S* was somewhat unresponsive, efforts were concentrated upon obtaining answers to the questions about with whom she ate or sat the previous week. The mean duration of interviews was six to seven minutes.

Prognosis. The ward psychiatrist, who had just completed a three-year residency training, independently made prognostic outcome ratings of *Ss* on the basis of personal impressions and case material within a month after the interviews. The following four-point scale was used:

1. *Potentially self-supporting.* Ss who can probably function at least marginally, and support themselves financially, outside an institution.
2. *Eventual release to custodial care, favorable.* *Ss* who are considered capable of adequate functioning in a protected environment under medical and psychiatric supervision outside a mental institution.
3. *Eventual release to custodial care, guarded.* *Ss* who currently are not stabilized, and who are hard to predict, but who, however, do seem to have some potential for improvement.
4. *Eventual release to custodial care, poor.* *Ss* who cannot possibly function outside a mental institution because of their mental status.

The same rater made two prognostic ratings on the scale for each *S*. A rating of actual prognosis (considering an *S*'s actually known resources outside the hospital), and a rating of ideal

prognosis (assuming that the same opportunities for support from human resources exist for each *S* outside the hospital.)

The internal consistency of the prognostic ratings was tested. After excluding the ratings of 1 *S* by random means, the ratings of 62 *S*s were divided into odd and even groups. No significant differences were found between the groups with respect to either actual or ideal prognostic ratings, the chi squares, corrected for continuity, $df = 1$, were .00 ($p > .99$) and .07 ($p > .70$), respectively. Neither were significant differences found in testing the first against the second half of the prognostic ratings ($p > .99$ for actual and $p > .20$ for ideal ratings).

Hospital adjustment. The level of hospital adjustment was defined in terms of the degree of responsibility required by any work that an *S* was doing in the hospital. The ward psychiatrist classified each *S* into one of five levels of hospital adjustment:

1. *S* leaves the ward to do non-ward work within the hospital.
2. *S* works on a ward, and is permitted, in the course of the work, to leave the ward unaccompanied, e.g., on errands.
3. *S* only works on a ward, her own or any other.
4. *S* does no work at all because of incapability.
5. *S* actively refuses to work despite being judged to be both physically and intellectually capable of working.

Naming[4]. An *S* with the ability to name accurately, on request, at least one of the patients with whom she regularly ate, was classified as a namer, and a *S* who did not have this ability was classified as a non-namer. A patient was taken to be accurately named by a *S* if enough of her first name, surname, or nickname was stated to allow a clear identification to be made of only one of the patients with whom *S* ate. The seating location of *S*s in the ward dining room had been determined by the staff for a duration of about two years. Each *S* was assigned a particular seat. All *S*s ate in groups of four, except for six *S*s who ate together.

Tension. Each *S* was rated on a tense/non-tense dichotomy during the interview. An *S* was rated as tense when at least one of any marked and unexpected changes of either her (a) body movement, (b) vocal tone, or (c) affective loading of language, was judged to attain a high level, and as non-tense when no such change was judged to attain a high level. Examples of behavior indicating a change attaining a high level are: (a) without warning, abruptly rising from a seat, or, without apparent purpose, moving a limb at least a moderate distance, (b) suddenly raising the pitch of the voice, (c) without warning, changing the topic of conversation and saying, *e.g.*: "I'm nervous and terribly unhappy."

All 63 *S*s were rated on tension by *E*, the observer independently rated 53 *S*s on tension. There was an inter-rater agreement of 84.9 per cent on the rating of 53 *S*s as tense or non-tense, the chi square, corrected for continuity, $df = 1$, was 24.45 ($p < .001$). Only the tension ratings of *E* were used in the analyses of data.

Table 1 shows the 63 *S*s classified into four groups: namer and tense, namer and non-tense, non-namer and tense, and non-namer and non-tense.

TABLE 1. DISTRIBUTION OF PROGNOSTIC RATINGS AND HOSPITAL ADJUSTMENT LEVELS OF PATIENTS GROUPED ACCORDING TO PERSON NAMING AND TENSION

Patient groups	N	Prognostic rating					Hospital adjustment level[c]				
		Actual[a]		Ideal[b]							
		4	3	4	3	2	5	4	3	2	1
Namer											
Tense	24	13	11	6	9	9	5	0	10	2	7
Non-tense	7	5	2	3	4	0	3	1	3	0	0
Non-namer											
Tense	13	13	0	9	4	0	2	8	3	0	0
Non-tense	19	18	1	10	9	0	4	9	6	0	0

[a]Actual prognostic ratings of one and two were not assigned.
[b]An ideal prognostic rating of one was not assigned.
[c]Defined as the degree of responsibility required by the work performed by an *S*.

[4]Naming was also defined as the ability to name accurately, on request, a patient, or patients, who, when free to choose, sat with the interviewee on the ward. "Sitting with" was defined as either at least one patient sitting immediately next to the interviewee, or one or more patients sitting near to the interviewee on at least two occasions. "Near to" was defined as sitting in the next chair but one or on a nearby facing chair. This definition did not sufficiently discriminate among *S*s, probably at least partly, because of the small number of observations that were made: one observation of the location of each *S* on each of five consecutive days.

The .05 level was set as an acceptable level of significance in all tests. The Fisher exact probability test[3] was used when expected frequencies were too small for the chi square test. All the chi squares with one *df* were corrected for continuity, and two-tailed tests were used throughout.

Results and Discussion

The namer[6] *Ss* had significantly better actual prognoses ($\chi^2 = 11.57$, $df = 1$, $p < .001$), ideal prognoses ($\chi^2 = 8.60$, $df = 1$, $p < .01$), and levels of hospital adjustment ($\chi^2 = 8.60$, $df = 1$, $p < .01$) than the non-namer *Ss*. In the test of ideal prognoses, favorable ratings were compared with the combined guarded and poor ratings, and in the test of hospital adjustment levels, the combined levels of non-ward work, and leaving the ward unaccompanied in the course of work, were compared with a combination of all the other levels. When the same comparisons were made with respect to the tense *Ss*, they were found to have significantly better ideal prognoses ($\chi^2 = 5.52$, $df = 1$, $p < .02$), and levels of hospital adjustment ($\chi^2 = 5.52$, $df = 1$, $p < .02$) than the non-tense *Ss*. Their actual prognoses were in the same direction but not at an acceptable level of significance ($p < .20$).

Further tests indicated that it was not naming or tension alone but their combination that tended to be associated with relatively better prognoses and hospital adjustment levels. Considering just the tense *Ss*, those who were namers had significantly better actual prognoses, ideal prognoses, and hospital adjustment levels than those who were not (Fisher exact $p = .006$, .02, and .02, respectively). Now taking only the namer *Ss*, those who were tense also had better (not at significant levels) actual prognoses, ideal prognoses, and hospital adjustment levels than those who were not (Fisher exact $p = .71$, .13, and .13, respectively). The following results indicate the potency of naming and tension combined: namer and tense *Ss* had significantly better actual prognoses ($\chi^2 = 10.39$, $df = 1$, $p < .01$), ideal prognoses ($\chi^2 = 18.05$, $df = 2$, $p < .001$), and hospital adjustment levels ($\chi^2 = 21.97$, $df = 2$, $p < .001$), than a group of all the other *Ss*. In the final test, three divisions of hospital adjustment levels were made: (a) *Ss* who did non-ward work, and *Ss* who left the ward unaccompanied in the course of work, (b) *Ss* who only did ward work, and (c) *Ss* who did no work at all, whatever the reason.

Further study of the namer and tense *Ss* showed that their prognostic ratings were significantly associated with their hospital adjustment levels (contingency coefficients, C = .51 and .48 for actual and ideal ratings, respectively, each $p < .001$, the upper limit of C is .71). The association is not surprising as in this hospital the responsibility taken by a patient in working mainly determines his prognosis. However, it was not expected that *Ss* would tend to be either namer and tense, or non-namer and non-tense. Naming and tension were, in fact, found to coexist significantly ($\chi^2 = 7.34$, $df = 1$, p. $< .01$). One could conjecture that patients characterized by naming are somewhat interested in, and open to, individuals. And, in consequence, such chronic patients may tend to become tense in relating with individuals.

To increase understanding of namer and tense *Ss*, variables were explored that might differentiate them from other *Ss*. These *Ss* were found to have significantly more accurate awareness of the patients who expressed a *liking* to eat or sit with them than a group of all other *Ss* ($\chi^2 = 14.14$, $df = 1$, $p < .001$). The criterion of accurate awareness was either (a) agreement between an *S*'s statement of who liked her and the patient's stated liking for *S*, or (b) an observation that a patient, who a *S* said liked her, sat with *S* on the ward when she was free to choose with whom she sat (the definition of "sitting with" was stated in Footnote 4). Thus, as well as being aware of the names of associates, namer and tense *Ss*, more than any others,

[6]To facilitate investigation it was deemed appropriate to report the extent of association of accurately naming at least one eating companion, with the accurate naming of any one patient who lived on the ward (in the interview *Ss* had been asked to name any of the patients who lived on the ward before they were asked to name the patients with whom they ate). The extent of association is expressed by a contingency coefficient of .52, $p < .001$ (the upper limit of the coefficient is .71).

can identify those of their associates who like them. Such an ability may well have good prognostic import in making more likely the forming of continuing relationships with others. Further tests showed no significant differences between the group of namer and tense *Ss* and a group of all other *Ss* with respect to age, level of educational attainment, duration of institutionalization, or duration of stay on the ward. In these tests, *Ss* who exceeded the median on a variable were compared with *Ss* who did not exceed the median. Neither did broad diagnosis significantly differentiate these groups. In this test, *Ss* with a diagnosis of schizophrenia (including 2 *Ss* diagnosed as paranoid state) were compared with all *Ss* with a diagnosis of either chronic brain syndrome or mental deficiency.

Summary and Conclusions

Indices of accurate person naming and tension were studied and found to relate to prognostic outcomes and hospital adjustment levels of 63 chronic mental patients from one ward. Compared to the others, the patients who were both tense and capable of accurately naming fellow-patients had better ratings of prognostic outcome, more responsible levels of hospital adjustment, more success in identifying fellow-patients who liked them, and similar broad diagnoses, ages, educational levels, durations of institutionalization, and durations of stay on the ward. Tension and the ability to name persons accurately were found to coexist in patients.

References

1. American Psychiatric Association. *Diagnostic and statistical manual. Mental disorders.* Washington: Author, 1952.
2. Peters, H. N. Traits related to improved adjustment of psychotics after lobotomy. *J. abnorm. soc. Psychol.*, 1947, *42*, 383-392.
3. Siegel, S. *Nonparametric statistics.* New York: McGraw-Hill, 1956.
4. Skottowe, I. Shock therapy; a plea for proportion in psychiatry. *Proc. roy. Soc. Med.*, 1939, *32*, 843-852.
5. Thomas, J. S. Some clinical examples of "dys-symbole": its relation to shock therapy. *J. ment. Sci.*, 1940, *86*, 100-108.

THE ATTITUDE OF INDUSTRIAL EMPLOYERS TOWARD HIRING OF FORMER STATE MENTAL HOSPITAL PATIENTS

VYTAUTAS J. BIELIAUSKAS AND HARVEY E. WOLFE
Xavier University *Longview (Ohio) State Hospital*

Introduction

Recently various governmental and private agencies have made special efforts to acquaint the general public with various aspects of mental illness. It appears that these efforts are beginning to pay dividends, because there is noticeably less stigma attached to the fact that one received psychiatric or psychological help. However, the situation seems to be different with those who have been patients in a state mental hospital. Apparently the public considers a state mental hospital patient (later referred to as SMHP) as one for whom there is little hope even though he may be treated and released.(Cf. 2, p. 14)

This negative attitude toward former SMHPs which is expected from the public by various workers in the field(Cf. 3) becomes a serious problem when such a former patient desires to become integrated into all aspects of our life, including gainful employment. Social workers, rehabilitation counselors, psychologists and psychiatrists have assumed that the attitude of the general public is shared by the employers and therefore many of them consider the finding of a suitable job one of the most difficult tasks for a former patient. This expected difficulty is accentuated by the widely accepted opinion among the professions concerned that vocational rehabilita-

tion of a mental patient is part of the therapeutic process.[7] First attempts to test empirically the validity of the opinion that the employers are not inclined to hire former mental patients date to as recently as 1958.[3, 4, 5]

The most extensive of these studies conducted by Olshansky et al.[5] indicated lack of "significant and substantial experience" with former mental patients among employers and therefore it did not obtain measurable data. Moreover investigations of Landy and Griffith[3] dealing with a small sample of employers in the metropolitan Boston area suggested a reverse trend to the generally accepted notion concerning the attitudes of employers toward former mental patients. According to these findings the greatest majority of employers interviewed expressed willingness to hire former mental patients. The present study further tested the suggestions of Landy and Griffith[3] on a better selected and larger sample of industrial employers of a different locale with attention given exclusively to SMHPs. A special attempt was made to compare the size of the employing company with the attitude toward hiring of former state hospital mental patients.

Method

Officials responsible for hiring in 251 manufacturing firms with 100 or more employees in a Midwestern city were interviewed as to their attitudes toward hiring of former SMHPs. These manufacturing firms were classified as "larger industries" by the local Chamber of Commerce and the sample included all such firms listed on the C. of C. roster of 1958. These industrial firms were employing from 100 to 14000 employees and they were classified by the C. of C. into 12 categories. An additional reclassification was accomplished by the authors in dividing all the industries into two categories: larger (L) and smaller (S). Both classifications are presented on Table 1.[1]

Table 1. Description of the Manufacturing Companies in Terms of Number of Employees, Number of Companies, Chamber of Commerce Code Letter and Designation by the Investigators

Number of employees	Number of companies in each category	C. C. code letter	Own designation of the companies	Number of companies in S and L category
100 - 249	136	D	S	
250 - 499	51	E	S	187
500 - 749	30	F	L	
750 - 999	8	G	L	
1000-1499	12	H	L	
1500-1999	5	I	L	
2000-2499	4	J	L	
2500-2999	1	K	L	
3000-3999	1	L	L	
4000-4999	0	M	L	
5000-5999	1	N	L	
Over 6000 to 14000	2	O	L	64
Totals 100-14000	251	D-O		251

Five female interviewers whose formal education included completion of undergraduate studies or more participated as aids in collecting of the data for this study. They all received a group orientation to the mechanics and goals of the project, industrial make-up of the community, vocational rehabilitation program at the hospital, population make-up at the hospital and the overall administrative procedure governing remission and employment of institutionalized patients. They contacted the employers individually by telephone using the same structured statement. While listening to the answer they made notes and on the basis of these notes they rated

[1]Companies which were leaving the area and did not care to give an answer were not included in these calculations.

the attitudes of the employers. A rating scale of 5 categories was employed, ranging from complete refusal to consider hiring former SMHPs to strong interest or even enthusiasm toward hiring of former SMHPs. All 251 firms were contacted and an answer was received from all of them. However, four firms were in process of moving and therefore their employing offices did not care to comment on the employment plans of their firms concerning former SHMPs. Therefore the answers of only 247 firms were used in most calculations. The data for the study were collected within a period of 90 days. A sample of 30 cases selected randomly from the cases of each interviewer was followed up by the co-author and rated independently. A standard statement was used again and the ratings of the co-author were compared with the ratings of the interviewers. This comparison yielded an $r = .94$. The findings were evaluated in terms of percentages. An attempt to examine each Chamber of Commerce category using the χ^2, one sample test had to be given up because in many cases the expected frequencies were smaller than 5 and in some instances even less than 1[1, 6]. Therefore in following the procedure suggested by Siegel([6, p. 42-47]) the categories were reduced to two: larger and smaller as indicated in Table 1. These two categories were then examined in using the χ^2 one sample test.

Results and Discussion

An evaluation of the findings in terms of percentages presented in Table 2 suggests that the majority of the employers interviewed were in favor of hiring of former

Table 2. Distribution of the Number of Answers and Percentages in Each Category According to the Size of the Company and Total Distribution

Code Letter	(1) Will not consider employing		(2) Hesitant to employing		(3) Slightly interested		(4) Interested		(5) Very interested		(6) Other	
D	8	6%	26	20%	40	29%	54	39%	6	4%	2	2%
E	4	8%	8	16%	22	43%	16	31%	1	2%		
F	5	17%	1	3%	9	30%	10	33%	3	10%	2	7%
G			1	12.5%	1	12.5%	4	50%	2	25%		
H	3	26%	2	16%	4	32%	3	26%				
I			2	40%			3	60%				
J	1	25%					3	75%				
K					1	100%						
J									1	100%		
M												
N					1	100%						
O					1	50%			1	50%		
Total	21	8%	40	16%	79	32%	94	37%	13	5%	4	2%

SMHPs, and that only 8 percent of them were completely against it. These results tend to substantiate previous findings[3, 5], and suggest that the general assumption concerning an overall negativistic attitude on the part of employers toward hiring of former SMHP's does not have empirical basis.

The evaluation of the findings in terms of χ^2 technique is presented in Table 3.

Table 3. Number of Answers in Each Category Given By Smaller and Larger Companies and the Corresponding Chi-Squares

Size of the Company	Number of answers in each category 1	2	3	4	5	Total	χ^2
Smaller	12	34	62	70	7	185	63.459
Larger	9	6	17	24	6	62	20.096
Total	21	40	79	94	13	247*	103.943

Df = 4 All χ^2 significant at .01 level

All chi-squares are significant beyond .01 level and suggest that the null hypothesis postulating no difference in rating preference among the employers cannot be accepted. The employers' preference shows definitely positive direction, and this is true regardless of the size of the company.

Summary and Conclusions

The employing officials of 251 manufacturing companies employing 100-14000 workers were interviewed as to their attitude toward hiring of former SMHPs. Their answers were rated by five interviewers on a rating scale of five. The findings were evaluated in terms of percentages and chi-squares. It was concluded on the basis of the findings that the employers interviewed in this study (a) tend to prefer certain attitudes toward hiring of former state mental hospital patients over the others, (b) this preference shows a highly positive attitude toward the hiring of former SMHPs, (c) only the minority of the employers object to hiring former SHMPs.

References

1. Cochran, W. G. Some methods of strengthening the common χ^2 tests. *Biometrics*, 1954, *10*, 417-451.
2. Coleman, J. C. *Abnormal psychology and modern life*. New York: Scott Foresman, 1956.
3. Landy, D., and Griffith, W. P. Employer receptivity toward hiring psychiatric patients. *Ment. Hygiene*, 1958, *3*, 383-390.
4. Landy, D., and Griffith, W. P. Placement of the emotionally handicapped—employer willingness and counselor practice. *J. Rehabilit.*, 1958, *24*, 17-18.
5. Olshansky, S., Grub, S. and Malamud, Irene. Employers attitudes and practices in the hiring of ex-mental patients. *Ment. Hygiene*, 1958, *3*, 391-401.
6. Siegel, S. *Nonparametric statistics for the behaviorial sciences*. New York: McGraw Hill, 1956.
7. Toppen, J. T. and Wolfe, H. E. Vocational rehabilitation as a therapeutic tool. *Dis. Nerv. System*, 1959, *9*, 425-426.

ATTITUDES OF RELEASED CHRONIC SCHIZOPHRENIC PATIENTS CONCERNING ILLNESS AND RECOVERY AS REVEALED BY A STRUCTURED POST-HOSPITAL INTERVIEW*

G. W. BROOKS, M. D. AND W. N. DEANE

Vermont State Hospital
Waterbury, Vermont

Introduction

In 1957, the Vermont State Hospital received a three-year grant to expand and study work in the rehabilitation of chronic schizophrenics. During the first full year of this research program, 100 patients were chosen for study who had been continuously hospitalized for an average of about 4 years each. Our results with this group of severely disabled patients have been quite encouraging. After a minimum of 6 months work with each member of the group, 67 were out of the hospital of whom 42 were employed full time in the community, 23 were in part-time employment or full-time training programs, and only 2 were unemployed.

A question relating to such hopeful results is, "How well are these people after they leave the hospital?" Their employment record suggests good functioning on at least a superficial level. Most of these patients, plus several others mostly chronic schizophrenics who had left the hospital in the past few years, were interviewed by the senior author during 1958 to gather information about the levels of their insight and views of their recovery process. The type of insight that was considered is that described by Bleuler[1], that is, the extent to which the patient speaks of his illness

*This study was made possible by a grant from the Office of Vocational Rehabilitation (OVR SP-180).

as something abnormal, something in the past that he has come out of or recovered from, and the extent to which he remembers the anguish of the situation with appropriate affect.

Procedure

The 78 patients included were 42 paranoid, 21 catatonic, 5 hebephrenic, 5 undifferentiated, 1 simple, 2 schizo-affective, and 2 non-schizophrenic. Patients had been released from a few days to 2 years (mean 4.6 months) at the time of interview. The questions presented in Fig. 1 were asked of each patient. The individual re-

Figure 1. Questionnaire for Post-Hospital Interview Rehabilitation Released Patients (Circle Relevant Word in Ratings)

Name.....Place.....Date.....Date of Release.....

1. "You appear to all of us to be much better. Do you feel that you are better?" Yes No
2. "Can you tell me what it was like before you started to feel better?" .....
   Rating:

	1	2	3	4
(a)	Clear, detailed answer	Hard to explain	Forgets or wants to forget	Vague or indefinite answer

3. "What helped you to get better?" .....
   Rating:

	1	2	3	4
(b)	A person Name.....	People or self	Gen'l atmosphere, work or drugs	"Time" or vague answer

4. "Was there any turning point when things started to clear up for you?" .....
   .....
   Rating:

	1	2	3	4	5
(c)	Definite dramatic point	Definite period of time	Vague or general	None	
(d)	Early in hospitalization ward	Disturbed ward	"Good ward"	Rehab ward	After release
(e)	A person Name.....	People or self	Gen'l atmosphere, work or drugs	"Time" or vague answer	

sponses to each of the four questions were written down verbatim or as closely as possible. The appropriate category which best described the response on each question was circled. In 17 cases (6 men and 11 women), the questionnaire was repeated. In these instances, the more complete questionnaire data were selected. For example, if on the first occasion only two of the four questions were answered but all were answered on the second interview, the later of the two questionnaires was selected. Also, if data were equal in amount on both questionnaires, we have selected the later one.

Interpretation of Data. For purposes of Chi Square analysis, the categories used for each question were dichotomized in the following ways. In question 2, "clear detailed answer" is thought of as clear and the other 3 categories are regarded as vague. Similarly in question 3, the categories "a person" and "people or self" are lumped as personalized, whereas "general atmosphere, work or drugs" and "time or vague answer" are lumped as impersonal.

Question 4 offers several foci of interpretation. In line c, we dichotomized 1 and 2 as dramatic or definite time and thought of 3 and 4 as vague time. Line d was analyzed from 3 points of view: first, the turning point was considered as occurring early in the hospital or late in the hospital with 1 standing for early and 2 - 4 standing for late. We did not include "after release" as a late response. We also compared rehabilitation ward (4) with non-rehabilitation ward (1, 2, 3, and 5). Lastly, we com-

pared turning point in hospital with turning point after release. This dichotomy set 1 - 4 off against 5. In line e, the personalized-impersonal dichotomy compared categories 1 and 2 against 3 and 4. The results are presented in Table 1. Not all released patients responded to all questions.

TABLE 1. DICHOTOMIZED SUMMARY OF POST-HOSPITAL INTERVIEWS

	Females	Males	Total
Present Psychiatric Status			
Out—Hospital	30	20	50
In—Hospital	9	8	17
In—Out Hospital	8	3	11
Total	47	31	78
Insight			
Clear	26	26	52
Vague	21	5	26
Total	47	31	78
Help			
Personalized	19	14	33
Impersonal	27	17	44
Total	46	31	77
Turning Point			
Dramatic or definite	38	29	67
Vague	7	2	9
Total	45	31	76
Rehab—Non-rehab			
Rehab	12	8	20
Non-rehab	31	23	54
Total	43	31	74
Early—Late			
Early	2	14	16
Late	41	17	58
Total	43	31	74
Hospital—After Hospital			
Hospital	35	28	63
After Hospital	8	3	11
Total	43	31	74
Personal—Impersonal			
Personalized	20	16	36
Impersonal	23	15	38
Total	43	31	74

RESULTS

Psychiatric Status. We first compared answers given to questions with psychiatric status. Dichotomizing the data from Table 1 into relapses and non-relapses, there was no significant difference between the two groups. How the questions were answered seemed to have no bearing on present status. We next dichotomized the data to compare those never having relapsed and those relapsed but returned to the community against those relapsed and in the hospital. In essence, successes at the present time were compared against failures at the present time. Again, there were no significant differences in response to questions on the part of these two groups.

Sex Differences. Men were more likely to have a clear insight into their illnesses than women. Twenty-one men expressed a clear insight whereas five were vague. Twenty-six women had a clear insight but twenty-one were vague, Chi $X^2 = 6.05$, $p = <.02$. Again, 14 men and 2 women associated the turning point of their illnesses with a

time early in hospitalization, while 17 men and 41 women felt the turning point came at a later point. This finding was highly significant, $X^2 = 20.08$, $p < .001$.

Inter-question Relationships. When the inter-relationships among questions were analyzed, the subjects with clear insight reported obtaining help from personalized and impersonal sources about equally, whereas persons with vague insight more often reported help from impersonal sources, $X^2 = 3.46$, $p < .07$. No other inter-question relationships of statistical or near statistical significance were observed.

Replies to Specific Questions. Table 1 demonstrates the following interesting information. Clear insight appears twice as often as vague insight (52 to 26). Dramatic and definite time is associated with the turning point more than 7 times as often as a vague sense of time concerning the turning point (67 to 9). The place at which the turning point came was in over 2½ cases to 1 not the rehabilitation ward (54 to 20). The turning point was late in hospitalization over 4½ times more often than early in hospitalization (58 to 16). The turning point came in the hospital almost 6 times as often as it came after leaving the hospital (63 to 11). The confidence intervals for the above data were determined by the method of Walker and Lev[2], using interpolations on a chart rather than exact computations, and the results are all significant beyond the .05 level.

Comparison by Diagnosis. Since the paranoids constituted the largest sub-category, they were compared with the remainder of the population on present psychiatric status and on answers to all questions but no significant differences were observed. Only one slight difference of non-statistical significance could be mentioned. Of 39 paranoids, 8 stated that the turning point came after hospitalization. Of 20 catatonics who answered this question, only 1 so stated. This difference approached the minimum .05 confidence limits.

Insight. On question 2 dealing with insight, all answers are considered and the clear-vague dichotomy used earlier is ignored. Since the population tended to speak of four major areas in describing its illness, all responses were classified as (a) affective state, (b) delusional system, (c) somatic complaints, and (d) behavioral problems. Some responses were classified in two or more areas, *e.g.*, if a man spoke of "spooks in his belly", he was obviously voicing a somatic complaint and at the same time talking about his delusional system. In such cases, we categorized the statement in two areas, the somatic and the delusional. Again, statements about "nerves", the term generally used by those giving a vague type answer, could mean almost anything. Unless this term was clearly amplified, it was viewed as bearing on affectivity and probably relating to anxiety. The data on insight in Table 2 indicate that the affective state was most often recalled in describing the illness. Often the terminol-

TABLE 2. CONTENT OF INSIGHT STATEMENTS POST-HOSPITAL INTERVIEWS QUESTION 2

Category		Number
Affective		58
Dissociation, depersonalization, disorientation, confusion, etc.	15	
Nerves or nervous	14	
Fear, anxiety, panic	12	
Nightmare or dream state	10	
Depression or despair	7	
Somatic Complaints: fatigue, weariness, headaches, etc.		22
Delusional System		21
Sense of persecution or feelings of hostility	13	
"Strange ideas"	5	
Voices	3	
Behavioral Problems: "Couldn't sit", "wanted to talk all the time", "overactive".		3

ogy used was highly dramatic, as for example: "Seemed as though everything left me, that I was in a new world. It was just like a nightmare only I couldn't wake up." The recounting of these experiences was often accompanied by some living of the experience again. Telling about it revived old anxieties but was often followed by relief.

Help. Patients reporting a personalized source of help cited staff members (22 cases), fellow patients (15 cases), and family members (3 cases) as sources of assistance. The way in which patients assist one another seems to be qualitatively different from the way in which staff or family members render assistance to patients. Table 3 helps to clarify this difference.

TABLE 3. TYPES OF HELPING RELATIONSHIPS REPORTED

Relationships Patients		Others		Sources of Relationships: Other Patients	Staff	Family
A	←	B	(gets help)	4	22	3
A	=	B	(mutuality)	8	0	0
A	→	B	(gives help)	3	0	0
Totals				15	22	3

Considerable patient-to-patient therapy goes on based either on a mutuality of relationships or one patient giving actual therapeutic aid to another. Three people said, in essence, "I got help by helping others." The pattern in any event is quite different from that involving patient and staff or patient and family wherein neither are seen as sources of help nor as groups with whom the patient may have a mutual relationship or to whom help may be given.

We were also interested to know, in the case of those released patients who talked about "general atmosphere, work or drugs", in what way the atmosphere was important and helpful. Little information could be gained at this point. Most who talked to this category spoke of work, routine, and medication. However, eight people amplified the term "atmosphere" by talking about a chance to rest or sleep. The answer seems important in the light of the several somatic responses to question 2 which frequently involved references to fatigue and weariness. It would seem reasonable that the chance to overcome these would be viewed as a source of help. However, only 2 people who had spoken of fatique in question 2 talked also of rest as the source of help.

Turning Point. Since many released patients had a dramatic or definite sense of time at which the turning point came, we attempted to ascertain what events were associated with this keen sense of time. It was impossible to break these answers down in detail since they were highly individualistic. In general, those who associated the turning point with a definite, though not dramatic, sense of time tended to talk about some pervasive and continuing change that came into their lives, such as a new hospital job, town privileges, the start of group therapy, permanent alterations in the pattern of family visits, etc. The event was of a continuous order eventually becoming incorporated as an expectation. The sense of time is associated with its beginning or acceptance by the individual as a continuing part of his life.

On the other hand, the people who spoke of dramatic time tended more often to cite a happening which was unexpected and occurred once and generally briefly but had the effect of altering the course of illness, such as the following: a patient was asked to sing in a play; another was given wanted clothes; one had an unexpected visitor who never came again and said, "We're all rooting for you"; still another asked for help and a doctor said, "I'll give you a pill; it might help"; someone asked for a cigarette and got it, etc. One patient, aware of this transient quality of the

experience, said, apparently with some mystification, "You wouldn't think a little thing like that would make so much difference."

There was another type of experience equally dramatic but totally ephemeral in which the patient said in effect, "I just got well one day." There was no memory of anything happening, but a change took place. Interestingly, these people were generally on a disturbed ward when the sense of turning point came. As far as they can recall, the day it happened was like any other day and the hour it happened no different than the hour before. Eight released patients recounted this kind of experience. Four were catatonic and four paranoid. However, when a catatonic described this type of experience, he did so with such words as, "I awoke", "I came alive", and so on. Indeed, at other points in the questionnaire, catatonics used this type of language. The paranoids expressed themselves a little less dramatically and were more inclined to talk about "things clearing up", etc. Perhaps the catatonic's mode of speech would be typical of a form of illness which so dramatically effects the bodily functions by affecting the will to movement. A diminution of symptoms perhaps could only be realistically explained by such somatically oriented terms as "I came to life" or "I woke up".

Summary

This study investigated attitudes towards mental illness and cure in a group of 78 chronic schizophrenics undergoing rehabilitation. Men possessed a clearer insight than women and saw their recovery as occurring early in hospitalization more frequently than women. Persons with clear insight saw the source of help as either personalized or impersonalized. Persons with vague insight more often saw the source of help as impersonal. This finding, however, falls slightly beyond the .05 minimum confidence limit. Most patients had a dramatic or definite sense of time associated with the turning point which was more often late than early in hospitalization. The turning point more often came during hospitalization than after it. Paranoid patients were more likely to see the turning point as taking place after hospitalization than catatonics, but this finding was not statistically significant. The affective state was most frequently recalled in descriptions of the illness.

Both staff and fellow patients were important therapeutically, with family members less so. Patients helped each other both by direct assistance and mutuality of relationship. Patients regarded the staff and family members as assisting only by actively giving help. No feelings of mutuality of relationship or assistance to staff or family members were apparent. The chance for rest and sleep may be an important part of hospital atmosphere.

Events associated with the time of turning point in the illness represented far reaching changes in patient activity or schedule or were dramatic, unexpected, transitory happenings leaving an indelible impression. Sometimes the patient had a sense of getting well to which great clarity of time attached, but did not include awareness of a particular happening. Catatonics sometimes described the dramatic sense of recovery without recall of particular events in somatic terms. The degree and quality of insight achieved by members of this chronic group suggest they they have experienced more than a superficial improvement.

References

1. Bleuler, Eugene. *Dementia Praecox or the Group of Schizophrenias.* New York: International Universities Press, 1950, p. 258.
2. Walker, Helen M. and Lev, Joseph. *Statistical Inference.* New York: Henry Holt, 1953, p. 461.

THE BENDER-GESTALT TEST AS A PREDICTOR OF LENGTH OF HOSPITALIZATION WITH MENTAL PATIENTS

DALE S. HIGBEE JOHN R. CLARKE WAYNE E. HENDERSON

Veterans Administration Hospital, Salisbury, North Carolina

Problem

Previous research [2, 4, 5] has demonstrated that the Bender-Gestalt may be of some value as a prognostic indicator in mental illness. The criterion used in these studies has been clinical judgment of improvement or alleviation of disturbing symptoms. With hospitalized mental patients this criterion may or may not be crucial, since it does not fully consider the more significant question as to how long the patient stays in the hospital and away from the outside community, or conversely, how soon he leaves the hospital and whether or not he later returns. The present study was undertaken to evaluate the Bender-Gestalt (B-G), as scored by the Pascal-Suttell method, as a predictor of length of hospitalization. Admittedly, this is not a simple variable, being determined by a variety of factors, but it is one of great importance and has the additional virtue of being readily quantified.

Method

Subjects. The Ss were 72 male veterans admitted to a Veterans Administration neuropsychiatric hospital and administered the B-G by the same psychologist in the course of individual diagnostic psychological evaluations. All subjects were examined during the period of 1 to 29 days following admission: mean number of days between admission and test was 8; median was 6; and only 6 cases were tested later than 15 days following admission. Other criteria for subjects were: 18 to 50 years of age on admission (median age was 34; mean age was 32.2 $\pm$ 6.7); dull normal or better intellectual functioning; functional diagnosis and absence of known organic brain damage; exclusion of all patients admitted for "Observation and examination" only.

Procedure. The Ss' records of hospitalization subsequent to the admission on which they were administered the B-G were checked, and the number of days they actually spent in a VA hospital over a two-year period was tabulated. The range was 2-730 days, the median being between 39 and 40. The Ss' copies of the B-G figures were independently rated by the two junior authors using the Pascal-Suttell method. The averages of the judges' ratings were used for the Bender scores, as their judgments were extremely reliable (r = .9749). The range of B-G scores was 6-193, the median being between 38 and 39.

Results and Discussion

Since the distribution of "days in hospital" was heavily skewed, nonparametric techniques of analysis were employed. Inspection of the data revealed no significant differences in number of hospital days between subjects with B-G raw scores above and below the median. The Mann-Whitney U Test also showed no significant difference at the quartile range, but a highly significant difference (P = .002) at the decile range. Thus, the B-G score does correlate significantly with length of hospitalization at the extremes, but only at the extremes.

A further analysis of our data revealed no significant relationship between B-G scores or number of in-hospital days and any of the following variables: age, education, presence or absence of acute symptomatology on admission. Highly related to length of stay in the hospital, however, is marital status. Regardless of their performance on the B-G, over a 2-year period following their admission, married Ss spent significantly less time in VA hospitals than Ss who were single, divorced or separated. One earlier study [4] reports similar findings in England, and research now in progress suggests that this holds true for hospitalized mental patients in both the VA [1] and state hospital systems[1] in the United States.

[1]Kramer, M. Personal communication.

SUMMARY

Seventy-two male veterans were administered the Bender-Gestalt (B-G) following admission to a VA neuropsychiatric hospital, and their records of hospitalization were followed over a two-year period. The B-G, scored by the Pascal-Suttell method, did not predict length of hospitalization except at the decile range, where it was highly significant. Age, education, and presence or absence of acute symptomatology on admission were unrelated to either B-G score or length of hospitalization. Marital status, however, proved to be related to length of hospitalization, married Ss spending significantly less time in hospitals over a two-year period than Ss who were single, divorced, or separated.

REFERENCES

1. JENKINS, R. L. and GUREL, L. Predictive factors in early release. *Mental Hospitals*, 1959, *10*, No. 9, 11-14.
2. LOTHROP, W. W. Relationship between Bender-Gestalt Test Scores and Medical Success with Duodenal Ulcer Patients. *Psychosom. Med.*, 1958, *20*, 30-32.
3. NORRIS, V. A statistical study of the influence of marriage on the hospital care of the mentally sick. *J. Ment. Sci.*, 1956, *102*, 467-486.
4. PASCAL, G. R. and SUTTELL, BARBARA. *The Bender-Gestalt Test: Quantification and Validity for Adults*. New York: Grune & Stratton, 1951, pp. 31-32.
5. SWENSEN, C. H. and PASCAL, G. R. A note on the Bender-Gestalt Test as a prognostic indicator in mental illness. *J. clin. Psychol.*, 1953, *9*, 398.

SPEED AND QUALITY OF PSYCHOMOTOR PERFORMANCE IN PSYCHIATRIC PATIENTS*

M. B. SHAPIRO, R. KESSELL AND A. E. MAXWELL

Institute of Psychiatry, Maudsley Hospital, London

INTRODUCTION

This paper investigates the effects of distraction procedure on the speed of psychomotor performance in co-operative psychiatric subjects. Slowness of psychomotor performance is a characteristic of co-operative psychiatric patients.[7] Foulds [3] found that dysthymic patients were speeded up on the Porteus Mazes by certain distraction procedures to a greater degree than were hysterics and psychopaths, implying that different mechanisms might underly the slowness of different diagnostic groups. It was therefore of interest to carry out a similar experiment on other nosological groups: normals, schizophrenics, and brain-damaged subjects.

On general and clinical grounds it was expected that brain-damaged subjects would be disorganized by interruptions (the so-called "catastrophe" effect), slowing down as compared with normals and schizophrenics. Precise expectations cannot be formulated concerning the differences between normals and schizophrenics. On some grounds they might be expected to speed up like the Foulds' dysthymics and on others they might be expected to slow down, though not as much as brain-damaged subjects.

---

*Acknowledgments must be made to the following physician superintendents for permission to work with their patients: Dr. A. Walk of Cane Hill Hospital, Coulsdon, Surrey; Dr. E. P. H. Charlton of Banstead Hospital, Sutton, Surrey; Dr. J. H. Watkin of Horton Hospital, Epsom, Surrey; and Dr. J. S. Harris of Claybury Hospital, Woodford, Essex. Also, to Miss J. le Prevost for her assistance with some of the computations, Mrs. M. Richardson who tested some of the subjects and scored their records, and the editor, Dr. F. C. Thorne, for his criticisms of the script. Part of the work reported in this paper was made possible by a grant from the Medical Research Council.

Test Battery

Porteus Mazes. This test was used as described by Foulds[3]. Each subject was presented with a series of 11 pencil and paper mazes, ranging in difficulty from the 5 year to the adult level. Two times were recorded. The first, starting time, was from the placing of the maze before the subject to the commencement of the movement of his pencil. The second, tracing time, was from the starting of the movement of the pencil to the end of the maze. Speed of performance was not referred to in the instructions. A split-second stop-watch, however, was present, and times were obviously being recorded. The distraction took the form of having the subject count after the experimenter, at the rate of one number per second, from the moment the maze was presented to him to the moment the maze was completed. The numbers were counted in their proper order.

Drawing Rotation Test. This test was given as a control for the effects of brain-damage. "Rotation" refers to the disorientation of the drawings of otherwise correct reproductions of abstract designs. Deviation from the correct orientation of each reproduction in the test was measured in degrees. Rotation has been shown to discriminate between brain-damaged and other patients [1, 9] and possibly to be affected by mechanisms associated with brain-damage [6].

Wechsler Sub-Tests. The Vocabulary, Similarities and Block Design tests were administered; the first to eliminate very dull subjects, and the others to amplify the description of the sample and to check some of the previous findings of Shapiro and Nelson [7] that performance on such tests was less affected by mental disorder than was psychomotor speed.

Babcock "Error-Free" Psychomotor Tests. These consisted of the time taken to (i) trace, with a pencil, two simple mazes, each shaped like a cross and with no blind alleys; (ii) write his name; (iii) write "United States of America"; (iv) write a simple sentence; (v) digit symbol; (vi) name colors; and (vii) name objects. These seven tests had been reported by Shapiro and Nelson [7] to be free of errors and to differentiate between psychiatric groups and normals more efficiently than more complex cognitive tests.

Subjects

There were 16 chronic brain-damaged, 16 chronic schizophrenics and 16 normals, each group being equally divided between males and females. The patients were in-patients of a number of hospitals mentioned in the acknowledgments.

Diagnosis. No schizophrenic patient with a history or symptomatology consistent with the possibility of brain-damage was included, *e.g.*, occurrence of head injury, periods of unconsciousness, leucotomy or E. C. T. Brain-damaged patients with complicating features were also excluded, *e.g.*, a report of a superimposed depression, and a history of schizophrenic withdrawal before the incurment of head injury.

Eleven of the brain-damaged subjects were G. P. I., one early Korsakoff, two post encephalitic Parkinsonians, one alcoholic, and one traumatic brain injury. Of the schizophrenics, eight were diagnosed unclassified schizophrenics, one a simple schizophrenic, four paranoid schizophrenics, and three hebephrenic schizophrenics.

The normal subjects were volunteers, mostly recruited from mental hospital staffs. All denied having been knocked unconscious, experienced fits or convulsions, abnormal birth, or having ever seen a doctor about a nervous complaint.

Psychomotor Speed. The two psychiatric groups were matched for psychomotor speed as measured by the total score on the eight Babcock "error free" tests. The results in Table 1 indicate that the matching was successful. It was not possible to match the normals, who turned out to be much faster than the other two groups. The F ratio for the three groups was 19.83, significant at the .001 level. Contrary to the finding of Shapiro and Nelson [7], most of our subjects made a few errors on these tests.

Total error scores were calculated for each subject. The mean was 4.5 for the organics, 2.0 for the schizophrenics, and 1.0 for the normals. An arbitrary cut-off point of 4.0 errors, dividing each group into two, produced a chi square of 15.42, significant at .001 level. The means and standard deviations of the Babcock "error-free" tests were recomputed with error not being taken into account (Table 1). The means and variances for the two methods of computation are almost exactly the same.

Co-operativeness. All subjects were rated for co-operativeness according to the scale of Huston et al [4]. Only three out of the 32 patients were given a C rating, defined as "needed some urging but did the task with perfunctory effort". The rest of the patients obtained ratings implying effective co-operation.

Chronicity. Patients with a history of less then three years' illness were excluded. The longest recorded illness was 24 years.

Age. The three groups were matched for age (Table 1), and ranged from 38 to 50 years.

Wechsler Vocabulary. No subject had a Wechsler Vocabulary Weighted Score of less than seven points. The organics and schizophrenics had nearly identical means and sigmas, and were of average ability compared with Wechsler's [8] young adults (Table 1). The normals had much higher scores, the F ratio being 8.45, significant at the .001 level.

TABLE 1. DESCRIPTIVE MEASURES

Measure	Normals (N = 16) Mean	S.D.	Brain-Damaged (N = 16) Mean	S.D.	Schizophrenics (N = 16) Mean	S.D.	F
Age	44.94	4.06	45.13	3.59	44.25	3.38	
Babcock (weighted for error)	15.72	2.20	10.45	2.70	10.59	2.90	19.83‡
Babcock (error omitted)	15.90	2.26	10.98	2.45	10.81	2.74	
Vocabulary (weighted score)	13.94	1.73	10.43	2.68	10.80	2.89	8.45†
Similarities (weighted score)	12.25	2.06	7.50	3.65	8.69	4.44	12.17‡
Block Design (weighted score)	11.19	2.74	7.06	3.82	8.63	3.38	5.72*
Combined weighted scores	37.38	5.72	25.19	8.74	28.13	9.43	9.42‡
Drawing Rotation (in degrees)	5.23	2.46	5.60	3.03	11.23	8.14	7.07*

‡P = .001
†P = .01
*P = .05

Wechsler Similarities and Block Design Tests. The schizophrenic and brain-damaged groups were not differentiated at an acceptable level on either of these tests, or on the sum of all three Wechsler tests, the t's being .944 for Similarities, 1.268 for the Block Design test and 1.005 for the sum of the three tests. All the eight t's for the differences between the brain-damaged and normals, and the schizophrenics and normals, reached the .005 level except one, which reached the .05 level. Details of this and all other computations are given by Kessell [3].

Two analyses of co-variance were performed to check the effects, on the Babcock psychomotor test scores, of biassed sampling from populations of different intelligence. The effects of the total of the three Wechsler weighted subtest scores were partialled out. The first analysis was carried out on Babcock scores calculated with error taken into account. The second analysis was carried out on Babcock scores calculated without taking error into account. The variance ratios were still significant at the .001 level, being 10.018 and 9.584 respectively.

Drawing Rotation Test. The brain-damaged subjects rotated more than the other groups. The F was 7.07, significant at the .01 level, after logarithmic transformations had been made to reduce heterogeneity of variance. The brain-damaged subjects are therefore probably typical with respect to psychomotor function.

PROCEDURE

All subjects were tested with the Porteus Mazes four times. For reasons not relevant to the purpose of this paper, the first trial of the first session of testing was administered under "ordinary" conditions and the second trial immediately following under "distraction" conditions.

A second session of testing was held after a lapse of between one and seven days. In this session half the subjects in each of the three nosological groups did the first trial under "ordinary" conditions, and the other half did the first trial under "distraction" conditions. The two conditions were strictly alternated for successive subjects. In the immediately following second trial of the second session, each subject was tested under the other condition. This was done to balance the influence of practice upon the two conditions.

RESULTS FOR STARTING TIME

Session 1. Analyses of variance were performed on the data in Table 2 after logarithmic transformations had been done to try to homogenize the variances. The "ordinary" trial and the "distraction" trial produced F's of 1.98 and 0.67 respectively, failing to reach the .05 level. The three groups tended to speed up significantly

and to an equal degree between conditions, F being 4.71, significant at the .05 level. The conditions-group interaction was not significant.

Session 2. A complex analysis of variance was performed. Only the difference between "distraction" and "ordinary" conditions reached the .05 level with an F of 4.17 (Table 2). The other main effects of diagnostic group, order, conditions and sex failed to reach the .05 level. These results are unequivocally opposite to those expected.

Results on Working Speed

Session 1. Rate, rather than time, was used as a measure, because the distance travelled by the pencil during the performance of a maze appeared to vary considerably from subject to subject, and from occasion to occasion for the same subject.

Table 2. Starting Time (in Seconds) and Tracing Rate (11 Trials of Inches per Second)

Groups	Session 1				Session 2			
	Ordinary		Distraction		Ordinary		Distraction	
	St.T.	Tr.R.	St.T.	Tr.R.	St.T.	Tr.R.	St.T.	Tr.R.
Brain-Damaged								
Mean	105.18	6.31	45.56	7.91	83.89	8.79	45.59	8.11
S.D.	71.24	2.74	14.67	3.00	67.49	6.01	27.28	3.83
Schizophrenics								
Mean	103.62	6.04	68.94	9.14	110.48	9.20	59.53	8.70
S.D.	69.81	3.04	97.18	3.82	229.76	3.84	72.03	3.28
Normals								
Mean	64.79	9.29	47.03	10.89	62.00	11.91	51.08	11.86
S.D.	35.38	2.52	41.38	3.02	62.98	4.81	34.25	3.06

A rate score for each maze was worked out in terms of inches per second, an optometer being used to measure distance travelled. Each subject's final score consisted of the sum of his eleven rate scores.

Table 3. Errors on the Mazes

Groups	Session 1		Session 2	
	Ordinary Conditions	Distraction Conditions	Ordinary Conditions	Distraction Conditions
Organics				
Mean	20.43	21.06	19.38	21.13
S.D.	2.78	5.98	6.50	5.98
Schizophrenics				
Mean	15.50	16.38	16.06	16.63
S.D.	4.47	3.40	3.84	4.53
Normals				
Mean	10.00	10.56	7.94	10.00
S.D.	6.10	5.45	5.23	7.05

The normals (Table 2) were faster than the other two groups, the F's being 4.20, significant at the .05 level. The *t*'s for the differences between the normals and schizophrenic group, and between the normals and organics were 3.337 and 3.108 respectively, both significant at the .01 level. The *t* for the difference between the psychiatric groups was 0.464, significant at the .20 level. The initial matching for psychomotor speed was therefore not unsuccessful.

Two analyses of co-variance were carried out. One partialled out the influence of the sum of all three Wechsler sub-tests, and the other partialled out the effects of the sum of Vocabulary and Similarities only, upon rate in the first trial. Both analyses produced an F which did not approach an acceptable level of significance, an outcome different from that reported above for the Babcock "error-free" tests.

Session 2. Each group produced almost identical means and standard deviations (Table 2) for both "ordinary" and "distraction" conditions, thus showing that distraction was having no effect on working rate. These results, for both sessions, are not in accordance with expectations.

Further Investigation of Rate

The effect of distraction on error. Distraction might have had no effect on rate, because its main effect was on error. Errors were therefore recorded for three kinds of response: (a) entry of a blind alley; (b) crossing of the line separating alleys, a relatively rare error; and (c) starting at the exit of a maze and working backwards to the entrance (a very rare error) occurring mainly in Maze No. 10 where entrance and exit were very close together. In the analysis of variance the differences between groups reached or exceeded the .001 level, the organics always being the worst. The differences between conditions, in both sessions, did not approach an acceptable level of significance.

The Effect of Unreliability. Kessell[6] gives a matrix of 45 product moment correlations between rate scores on 10 successive administrations of the Porteus Mazes to 16 subjects: eight of our female schizophrenics and eight of our female organics. The correlations range from .71 to .98; the mean being .87. Unreliability cannot be the cause of the unresponsiveness of rate to distraction.

Discussion

The Relation of Slowness to Psychiatric Disorder. This study appears to confirm the previous finding that the Babcock "error-free" psychomotor speed tests are more sensitive to mental disorder than are other tests. This sensitivity appears related to the aspect and type of task, and the conditions of its performance. This is evidenced by the fact that the Porteus Mazes, unlike the Babcock Speed tests, failed to differentiate between the groups, when the influence of the Wechsler tests was partialled out.

The Porteus Mazes differ from the Babcock tests in: (a) involving problem solving and (b) not being explicitly speeded in the instructions. Relevant is the finding of Shapiro and Nelson[13] that Furneaux's test of speed of verbal reasoning did not differentiate at an acceptable level between nosological groups under non-speeded conditions, though it did so under speeded conditions. Relevant too, are the facts that, unlike working time, starting time did not differentiate significantly between groups and was significantly shortened by distraction.

The Relation between Speed and Quality of Performance. The brain-damaged group was inferior to the other two groups on three separate measures of quality of psychomotor performance: (a) errors on the Babcock tests; (b) errors on the Porteus Mazes; and (c) the rotation effect. One can rule out intelligence and age as causes, as the brain damaged and schizophrenic subjects were quite well matched on these two variables. One can also rule out conditions of testing, as the Babcock tests were speeded, while the Porteus Maze and rotation tests were not. It would seem that quality of psychomotor performance and speed of psychomotor performance have different relationships to mental disorder, and therefore possibly have different underlying mechanisms. This generalization does not necessarily hold for errors of intellectual function as none of the Wechsler tests differentiated significantly between organics and schizophrenics.

Response to Distraction. The similarity of our three groups in response to distraction is confirmed by a similar finding by Campbell[2] for depressed, schizophrenic and

normal subjects. Two different, but not mutually exclusive, interpretations are possible. First, the mechanisms underlying speed of performance might be similar in a large number of nosological groups. Taking Foulds' results into account, we might conclude that it is only hysterics and psychopaths who differ from other subjects, the difference possibly being due to a greater distractibility in hysterics and psychopaths in the first place. Second, starting time and working rate on the Porteus Mazes might not be the most appropriate measures. Our results might have been different if, instead, we had used the "error-free" Babcock tests.

Further advance in our understanding of psychomotor slowness in co-operative psychiatric patients is dependent on establishing exactly the conditions under which such slowness is manifested, and how this slowness is related to quality of psychomotor performance. The results of the present study give some indication of the kind of conditions which are relevant.

Summary

A battery of psychomotor and cognitive tests was given to 16 normals, 16 co-operative schizophrenics and 16 co-operative brain-damaged subjects. The main conclusions are: The psychiatric patients were slowest in tasks which did not involve problem solving and which emphasized speed in the instructions. Speed and error of psychomotor performance differ from each other in their relation to psychiatric groups and hence, possibly in their underlying mechanisms. The speed of psychomotor performance was affected in a similar manner by "distraction" in the three groups, and hence the mechanisms mediating psychomotor speed might be similar for all three groups.

References

1. Babcock, H., and Levy, L. *Test and manual of direction, the revised examination for measurement of efficiency of mental functioning.* Chicago: C. H. Stoelting, 1940.
2. Campbell, D. *A study of some sensory-motor functions in psychiatric patients.* Unpublished Ph.D. Thesis, University of London, 1957.
3. Foulds, G. A. Temperamental difference in maze performance. Part II. The effect of distraction and electro-convulsive therapy on psychomotor retardation. *Brit. J. Psychol.*, 1952, *43*, 33-41.
4. Huston, P. E., Shakow, D., and Riggs, L. A. Studies of motor function in schizophrenia. II. Reaction Time. *J. gen. Psychol.*, 1937, *16*, 39-82.
5. Kessell, R. *An investigation of the factors affecting speed of response in psychiatric patients with special reference to distraction.* Unpublished Ph.D. Thesis, University of London, 1951.
6. Shapiro, M. B. An experimental investigation of the block design rotation effect. *Brit. J. Med. Psychol.*, *27*, 1954, 84-88.
7. Shapiro, M. B., and Nelson, E. H. An investigation of the nature of cognitive impairment in co-operative psychiatric patients. *Brit. J. Med. Psychol.*, 1955, *28*, 239-256.
8. Wechsler, D. *The measurement of adult intelligence.* Baltimore: Williams & Wilkins Co. 1944.
9. Yates, A. J. The rotation of drawings by brain-damaged patients. *J. abn. soc. Psychol.*, 1956, *53*, 178-181.

DISTANCE CONSTANCY IN SCHIZOPHRENIC AND NON-SCHIZOPHRENIC MENTAL PATIENTS

T. E. WECKOWICZ AND R. HALL[1]

Saskatchewan Hospital, Weyburn (Sask.) Canada

PROBLEM

In a previous study[7], schizophrenic patients were shown to have poorer distance constancy than normal controls. This is in keeping with the poor size constancy in schizophrenic patients reported by Weckowicz[6] and Crooke[1]. It will be remembered that a standard physical unit of distance, one yard for instance, subtends progressively smaller visual angles with the increased viewing distance. Conversely, the same visual angle is subtended by distance units of increasing length with increased viewing distance. Distance constancy is the mechanism responsible for the ability to make estimations in accordance with viewing distance, irrespective of the visual angles. Therefore the subjects who have good distance constancy will give estimations approximating physical distance units independent of the viewing distance; while the subjects who have poor distance constancy will give estimations greater than physical distance units and the lengths of these estimations will increase with the viewing distance, according to the visual angle.

In the previous study schizophrenic subjects and normal controls were asked to judge successive one-yard intervals in depth towards the horizon by telling the experimenter to stop a marker which was being moved away from another stationary marker. The point to which the first marker had been moved before the distance was judged to be one yard became the starting point of another "one-yard" judgment. It was found that the measurements of these subjective one-yard intervals formed a gradient of intervals of increasing length. This gradient was much steeper in the schizophrenic patients than in the normal controls which indicated that distance constancy was poorer than in normals. Moreover it was found that in the schizophrenic subjects there was a positive correlation between size and distance constancies. Those with poor distance constancy had also poor size constancy. This correlation was not unexpected, since distance can be considered as size in the horizontal plane.

Before poor distance constancy could be attributed to schizophrenic illness *per se* it is necessary to find out whether schizophrenic patients differed from non-schizophrenic patients. Otherwise it could be argued that poor distance constancy was not due to schizophrenic process, but to the non-specific effects of mental illness and hospitalization. Moreover in the previous experiment the distances from which the judgments were made were not controlled directly but only subsequently in the course of statistical analysis of the data. This paper reports an experiment in which the distance constancy of schizophrenic patients was compared with that of non-schizophrenic patients. The technique of the experiment was modified to make it more rigorous.

SUBJECTS

A completely new sample of 20 schizophrenic patients was used in whom the diagnosis of schizophrenia had been firmly established by all the psychiatrists who examined them. It consisted of 9 females and 11 males, all chronic patients, but most of whom had freedom of the grounds. There were 8 hebephrenic, 4 paranoid, 2 simple and 3 undifferentiated schizophrenics. The mean age was 39.6 years (S. D. = 7.0). The mean length of stay in the hospital was 10 years (S. D. = 7.3). No patient in whose case there was a possibility of mental deficiency or organic brain disease was included. No patient undergoing insulin or electric shock treatment was included. Only the patients in whom a satisfactory cooperation during the experiment could be obtained were included in the sample.

The non-schizophrenic sample consisted of 9 females and 10 males, including 5 depressions, one excited manic-depressive psychosis, two cases of epilepics, 5 psychopathic personalities,

[1]This work was carried out using funds provided by the Rockefeller Foundation and a National Health grant (Ottawa). We would like to express our thanks to H. Osmond, A. Hoffer, R. Sommer and Miss G. Witney.

2 alcoholics, one hypochondriacal reaction and one immature personality. The mean age of non-schizophrenic patients was 39.5 years (S. D. = 14.9). The mean length of stay in the hospital was 6 months (S. D. = 1.5 years). No patient with a possibility of mental deficiency or organic brain disease was included. The social, cultural and educational backgrounds in both samples were very similar. No patient with gross refraction error, unless corrected by glasses, was included in either sample. The samples could not be matched as to the length of hospitalization, since there were very few non-schizophrenic long-stay patients (excluding defectives and organic patients) in the hospital. However, all the schizophrenic patients were on the grounds frequently and hence were even more familiar than the other group with the field where the experiment took place.

Procedure

An outdoor experiment was carried out on a flat, open prairie with an open horizon. A small cart with a white triangular marker was moved by a series of strings and pulleys away or towards the subject. It was held on a straight path by a white nylon rope stretched from an iron peg at the subject's feet to another peg 200 yards away. A red painted peg was used as the stationary marker. The strings and pulleys were operated by an assistant who stood behind the subject.

In the previous experiment the subjects had to double and divide into halves standard distances before judging one-yard intervals. As there was no statistically significant difference between the normals and the schizophrenics in this task, only the judging of one-yard intervals was used in the present experiment. Also, in the previous experiment the successive subjective yards were estimated by different subjects at different distances because the subjects who overestimated the first few yards estimated subsequent yards from a greater distance than subjects who made shorter estimations on the first trials. This possible confounding of individual yard estimations with the distance from which these estimations were made was to some extent rectified in the subsequent analysis of the data by comparing the yard estimations by the normals and the schizophrenics made at approximately the same distance.

In the present experiment it was decided to compare the judgments of one-yard intervals made by the two groups at the same standard distances. In the previous experiment it was found that the difference between schizophrenics and normals in one-yard judgments became significant only after the distance from which the judgments were made was more than 10 yards. Therefore in the present experiment it was decided to use 10 standard distances ranging from 11 yards to 20 yards. In order to reduce the possibility that non-visual cues would be used by the subjects the order of presentation of these distances was randomized. The subjects were divided into pairs, each consisting of one schizophrenic patient and one non-schizophrenic patient. For both subjects in each pair the order of presentation was the same, but it was different for different pairs.

The following instructions were given to the subject: "We are going to move the cart with the white triangle away from the red peg. Would you tell us to stop when the near side of the cart is one yard from the red peg?" After the subject said, "Stop the cart," he was told to turn his back. The distance the cart had travelled was measured. The red peg was then removed and driven into the ground at the distance which was next in the order of presentation. This procedure was repeated ten times.

Results

Because of the obvious lack of homogeneity of variance, non-parametric statistical methods were used in the analysis of the data. In both groups the objective lengths of the subjective yards had a tendency to increase with the distance away from the subject. The differences between the estimations of the two groups were tested by the Mann-Whitney U Test. The results in Table 1 show that at all the

Table 1. Estimations of One-Yard Intervals

Distances in yards	Medians in feet			
	Schizophrenics	Non-schizophrenics	U	p
11	7.1	5.2	124	$< .05$
12	7.5	5.1	101	$< .01$
13	7.9	5.2	126.5	$< .05$
14	7.7	5.6	105	$< .01$
15	8.5	5.5	93.5	$< .01$
16	8.9	5.4	77.5	$< .001$
17	8.0	5.7	128.5	$< .05$
18	8.3	5.7	125	$< .05$
19	8.2	5.7	107	.01
20	9.7	5.7	81	$< .001$

*One-tail test.

distances the median estimations of the schizophrenic patients are longer than the median estimations of the non-schizophrenic patients. The medians also increase with the distance more rapidly in the schizophrenic group than in the non-schizophrenic group. The differences between the two groups are statistically significant at all ten distances. These results show that in schizophrenic patients judgments of one-yard intervals in depth are determined to a greater extent by the visual angles than the same judgments in non-schizophrenic patients.

Thus distance constancy in schizophrenic patients is poorer than distance constancy in non-schizophrenic (non-organic) patients. However before this conclusion is warranted another possible explanation has to be considered. It is well known that reaction time of schizophrenic patients is slow. There is a possibility that the difference between the objective lengths of the subjective yards in both groups is due to the fact that there is a delay in the reactions of schizophrenic patients and the cart is pulled away a longer distance before they say "stop," than is the case with the other group.

To test this possibility another experiment was carried out. The technique was modified in the following way: the cart, instead of being pulled away from the red peg, was pulled towards the red peg from a constant distance of 50 yards. (For instance when the red peg was set at the distance of 11 yards, the cart was pulled towards it from the distance of 61 yards). If the reaction time was the variable responsible for the difference it should influence the results in the opposite direction when the cart is being pulled towards the red peg. Five schizophrenic patients who estimated long intervals as one yard and five non-schizophrenic patients who estimated short intervals as one yard were chosen. They were given the instruction, "We are going to move the cart with the white triangle towards the red peg. Would you tell us to stop when the approaching cart is one yard away from the red peg?" The same order of presentation of distances for each patient and the same procedure as on the previous occasion was used. The lengths judged as subjective yards tended to be *longer* in both groups when the cart was pulled towards the red peg, than on the occasion when it was pulled away from the red peg. For nine distances (from 12-20 yards) the five longest judgments of one subjective yard were made by the five schizophrenic patients and the five shortest judgments by the five non-schizophrenic patients. For the distance 11 yards: 4 schizophrenic patients and one non-schizophrenic patient made the five longest judgments of one subjective yard; 4 non-schizophrenic patients and one schizophrenic patient made the five shortest judgments of one subjective yard. Thus, almost all of the schizophrenic patients' judgments were longer than those of non-schizophrenic patients. Therefore it can be seen that the change in the direction of the movement of the cart did not make the slightest difference as far as the performance of both groups in relation to each other was concerned.

TABLE 2. CORRELATION BETWEEN PULLING CARD TOWARDS AND AWAY FROM PEG

Distance in yards	*rho*
11	.45
12	.70*
13	.83**
14	.75**
15	.71*
16	.88**
17	.81**
18	.89**
19	.82**
20	.79**

*$p < .05$
**$p < .01$

The Spearman Rank Difference Correlations for the ten subjects (5 schizophrenics and 5 non-schizophrenics) between their performances when the cart was pulled away from the red peg and when it was pulled towards the red peg are presented in Table 2. Nine correlations out of ten are statistically significant. It can be concluded that the differences between the objective lengths of the subjective one yard judgments in schizophrenic and non-schizophrenic patients are not due to the possible difference in the reaction time and also that the results are to a great extent reproducible even when the technique is changed, with the cart being pulled away from the red peg instead of being pulled towards the red peg.

Discussion

The results of this experiment are in full agreement with the results of our previous study[6] of distance constancy in schizophrenics which established that distance constancy is poorer than in normal controls. This study which used another sample of schizophrenic patients, established that distance constancy is poorer in schizophrenic patients than in non-schizophrenic (non-organic) patients. Thus it can be assumed that, with a possible exception of some patients with organic damage to the brain, poor distance constancy is quite specific for schizophrenic illness[6].

This is in complete agreement with the previous study[6] of one of us, which showed that the size constancy in schizophrenic patients is poorer than in normal controls and non-schizophrenic patients.

The implications of these findings for the phenomenological study of the subjective world of the schizophrenic were discussed in the previous paper[7]. It was pointed out that schizophrenics lived in a "flatter" world than non-schizophrenics, that their world was lacking in depth or three-dimensionality of the world of normal people. Here we would like to mention some possible mechanisms, responsible for the impaired depth perception of schizophrenic patients. In the previous study it was found that the sense of depth becomes significantly poorer in schizophrenic patients than in normal controls only beyond ten yard distance. This suggests that the distance cue produced by lenticular accommodation can be ruled out. Also the distance cue produced by the ocular convergence can be ruled out to a great extent, although some authorities, such as Graham[3], give 20 yards as the limit of the effectiveness of the ocular convergence in supplying depth cures. It leaves relative size, interposition, linear perspective, aerial perspective, monocular movement parallax, light and shade effects as monocular cues and stereoscopic vision as a binocular cue. Of these, probably stereoscopic vision, monocular movement parallax and resolution of texture gradients in linear perspective, the importance of which was stressed by Gibson[4], are the most important. Eysenck, Granger and Brengelmann[2] studied binocular vision in normals, neurotics and psychotics with inconclusive results but a definite trend for both psychotics and neurotics to give poorer performance than normals. In one of the tests (Worth Four-Dots Test, assessing binocular vision as "distance") the difference is statistically significant. They did not describe their psychotic sample, but it can be assumed that large proportions were schizophrenics. In their study of binocular vision they used distances shorter than ten yards. There is ample evidence[3] that stereoscopic vision acts as a space cue through a considerable range of distance. Matsubayashi[5] has shown that distance of observation with its associated condition of convergence influences the acuteness of stereoscopic vision, It is important to study stereoscopic vision in various psychiatric groups at distances longer than 10 or even 20 yards where it is free from the convergence effect. Apart from the stereoscopic vision the texture gradients described by Gibson may be important. If this is true, it is possible that schizophrenic patients, due to either a peripheral or a central functional defect are not capable of using all the information about space contained in the texture gradients. Monocular movement parallax may be another important factor. Non-schizophrenics may make a greater use of the monocular movement parallax cues given by small lateral head movements when observing the receding cart. Their attitude towards the per-

ceived object may be more "active" than that of schizophrenics and while observing an object they may make more often small shifts in the line of regard.[2]

Summary

Distance constancy was studied in schizophrenic and non-schizophrenic patients. Distance constancy was poorer in the schizophrenic than in the non-schizophrenic patients. The mechanisms possibly responsible for poor distance constancy in schizophrenic patients are discussed.

[2]In our experiment gross changes in line of regard were prevented by the instruction which forbade Ss to change their standing place. In future experiments even finer changes in the line of regard could be prevented by using a chin support.

References

1. Crooke, T. G. Size constancy and literalness in the Rorschach test. *Brit. J. Med. Psychol.*, 1957, *30*, 99-106.
2. Eysenck, H. J., Granger, G. W. and Brengelmann, J. C. *Perceptual processes and mental illness.* (Maudsley Mong. No. 2). London: Chapman & Hall, 1957.
3. Graham, C. H. "Visual perception" in *Handbook of experimental psychology.* (S. S. Stevens, Ed.), N. Y.: John Wiley, 1951.
4. Gibson, J. *The perception of the visual world.* Boston: Houghton Mifflin, 1950.
5. Matsubashi, A. Forschung uber die Tiefenwahrnehmung. *Act. Soc.* III, *Opthalm. Jap.*, *42*, 480-491.
6. Weckowicz, T. E. Size constancy in schizophrenic patients. *J. ment. Sci.*, 1957, *103*, 475-486.
7. Weckowicz, T. E., Sommer, R. and Hall, R. Distance constancy in schizophrenic patients. *J. ment. Sci.*, 1958, *104*, 1174-1182.

A COMPARISON OF THE BENDER-GESTALT AND FLICKER FUSION AS INDICATORS OF CENTRAL NERVOUS SYSTEM INVOLVEMENT[1, 2]

FREDERICK L. MC GUIRE

U. S. Navy Medical Neuropsychiatric Research Unit
San Diego, California

Introduction

One of the most difficult diagnostic problems in clinical psychology and psychiatry is to decide whether a given patient has a disorder of the central nervous system. If the presence of an organic factor is not obvious to the physician, the psychologist is often asked to give an opinion about such damage as well as a description of personality. Two methods often used to try to determine the presence of brain lesions are the Bender Visual Motor Gestalt Test (1) and Flicker Fusion Frequency (2). The purpose of this paper is to present a comparison of these two methods as applied to the detection of CNS pathology.

Procedure

The subjects consisted of two groups of 21 men each matched for age. The control group consisted of healthy males known to be free of any symptoms indicating damage to the central nervous system or of visual defects other than refractive error correctable by glasses. Fifteen were Navy hospital corpsmen, one a physician, and four patients at a U. S. Naval Hospital. The latter four were retired military personnel (ages 62 through 65) who were hospitalized for other than central nervous system disorders, such as arthritis or a broken limb. These four subjects were included in

[1]The author wishes to express his gratitude for the cooperation and professional skill rendered by each of the following five judges: Dr. Max Hutt, Dr. Lauretta Bender, Dr. Fred Wissner, Dr. Gerald Pascal, and Dr. Gerald Briskin.

[2]The views expressed in this publication are not necessarily those of the Department of the Navy.

the control group only because they demonstrated a high level of vigor and alertness, and were leading active lives just prior to their admission to the hospital.

The experimental group were hospitalized patients which had been processed by the hospital staff and given a final diagnosis. Each had also been examined and declared to be free of defects of the visual globe (including fundi), except refractive error which was correctable by glasses. The type of central nervous system pathology represented by these subjects included demonstrable brain lesions, cerebral vascular accidents, demonstrable tumors, and multiple sclerosis.

The light source for the flicker was provided by a 50 watt, 6 volt bulb on direct current and at a brightness level of 18.9 foot lamberts. The light beam was interrupted by a rotating disc with a pie-shaped wedge cut from it providing a light-dark ratio of 50-50. The visual angle was 1 degree. A more detailed diagram of the apparatus and its use may be found in McGuire[3]. The beginning speed of the disc was 9.0 flashes per second in the case of the CNS group and 30 flashes per second when presented to the normal subjects. The stimuli were presented in ascending steps of 1.5 fps with an exposure time of two seconds.

The same 42 patients were also administered the Bender-Gestalt Test immediately following the Flicker Test. Since the "scoring" of this test is non-objective, the 42 tests were photostated and distributed to five judges, each an authority on the Bender-Gestalt.[1] Each judge was asked to decide, on the basis of the drawings, whether or not each subject was (a) not "organic", (b) maybe "organic", or (c) "organic". In addition, each judge was asked to list for each separate figure and for the test as a whole those "signs" which contributed to his diagnosis.

Results and Discussion

The two methods were compared on their ability to correctly label each of the 42 subjects. In the case of flicker, the resulting distributions of fusion scores were compared and a cut-off score noted which would allow for a maximum number of experimental subjects to be listed below a certain score and a maximum number of control subjects above that score. The cut-off score in this case was placed at 57.2 fps, that is, all persons scoring above this point were labeled as normal, and all those scoring at or below this score were labeled as having CNS involvement. For the normals the flicker scores ranged from 38.5 fps to 93.7 fps with a mean of 60.7 fps. Fifteen, or 71% of these subjects fell above the cut-off score. For the CNS subjects the flicker scores ranged from 9.0 fps to 72.7 fps with a mean of 48.0 fps. Fifteen, or 71% of this group fell below the cut-off score.

The five Bender-Gestalt judges did not agree unanimously in their diagnoses. Thus, different rates of success of diagnosis were established by use of different criteria: (a) when three of the five judges agreed, (b) when four of the five judges agreed, and (c) when all five judges agreed. Each of the "Maybe" responses was listed as a positive diagnosis in order to facilitate analysis. This technique produced insignificant changes in the data.

Flicker correctly labeled 30 of the subjects while 29 of them were correctly labeled by the use of the B-G when three of the judges agreed, 21 when 4 judges agreed, and 16 when all five judges agreed. The difference between the number of successful diagnoses made by Flicker and the Bender-Gestalt does not approach the .10 level of confidence except when 5 judges agreed ($\chi^2 = 4.38$; $P = .04$).

It is important to note that flicker produced a rate of selection that was equal to or better than that of the B-G, and did so on the basis of a single, objective measurement which can be obtained by personnel with no clinical training. However, it was found that those subjects which were correctly identified by each method were not necessarily the same individuals. A contingency coefficient of .21 between Flicker and B-G resulted when agreement of 3 judges was used as a criterion, and lesser coefficients obtained for the other two B-G criteria. This correlation is not statistically significant ($P = .30$), and suggests that the two tests measure different things.

It would appear that there is some gain in accuracy when the two techniques are used in combination. However, such a gain would exact the penalty of leaving several cases undiagnosed, that is, the tests will disagree. For example, for each of the respective B-G criteria the two tests correctly labeled 88%, 86%, and 93% of of those cases about whom they agreed; but they agreed only 60%, 50% and 36% of the time. This, of course, is further demonstration of the fact that the two tests are not highly correlated.

Summary and Conclusions

The Bender-Gestalt Test and a flicker fusion test were compared according to their ability to correctly distinguish between a group of normals and a group of patients with CNS involvement. The B-G tests were submitted to five judges, each of whom made a decision as to whether or not the test indicated the presence of CNS pathology. Three different criteria were used to decide if a given individual was labeled as "organic" or "not organic" by this test: (a) when 3 of the 5 judges agreed, (b) when 4 of 5 judges agreed, and (c) when 5 of 5 judges agreed. FFF correctly identified 71% of the forty-two subjects. For each of the three criteria of judgment, the B-G test was able to correctly identify 69%, 50% and 38% of the subjects, respectively. The correlation obtained between flicker and the Bender-Gestalt was not statistically significant, suggesting that they measure different things.

Flicker and the B-G test agreed on a diagnosis 60%, 50% and 36% of the time for each of the three B-G criteria, respectively. When the two tests are in agreement the accuracy of diagnosis may improve, but the total number of cases about whom a diagnosis is possible drops sharply.

References

1. Bender, Lauretta. A Visual Motor Gestalt Test and Its Clinical Use. *Research Monographs No. 3*, The American Orthopsychiatric Association, New York, 1938.
2. Landis, Carney. Determinants of the CFF Threshold. *Physiol. Rev.*, 1954, *34*, 259-288.
3. McGuire, Frederick L. A Comparison of Two Kinds of Flicker Phenomena As Indicators of CNS Involvement. *J. nerv. ment. Dis.*, 1958, *126*, 540-546.

COMPARISON OF REPRODUCTION AND RECALL SIZE DEVIATIONS IN THE BENDER-GESTALT AS MEASURES OF ANXIETY

DANIEL GAVALES AND THEODORE MILLON

Lehigh University

On the Bender-Gestalt, one of the principal indicators of anxiety is an overall decrease in the size of reproduction [1, 4]. Unfortunately, the validity of this particular measure has not been adequately checked. Several investigators have also hypothesized that the Bender recall phase would reflect situational and personality variables better than the initial reproductions [3, 4, 7]. The logic for this suggestion follows from a basic assumption in projective methodology: the less structured a procedure, the easier one will detect the personal qualities of the perceiver. Data with regard to the special usefulness of the Bender recall technique has been contradictory. It was the purpose of the present study, first, to check on the validity of the hypothesis that anxiety would be tapped by decrements in the size of the Bender-Gestalt stimuli, and, secondly, that the recall phase would be more sensitive to differences among Ss in this variable than the reproduction phase.

Method

Subjects. One hundred and ninety five undergraduates were administered the Taylor Manifest Anxiety Scale [6] with the highest 40 (Taylor-High) and lowest 40 (Taylor-

Low) selected as experimental Ss. On a random basis, 20 Taylor-High and 20 Taylor-Low scorers performed the Bender in an anxiety producing atmosphere while the remining Ss took the test under nonstress conditions.

Experimental Conditions. In the *induced anxiety group* (IA) Ss were subjected to a series of repeated failures with several pseudo-intellectual tasks immediately prior to the administration of the Bender. Following the standard administration [1] of the initial reproduction phase of the test, a one-minute time delay ensued during which miscellaneous data relevant to academic interest and standing were collected. After this delay the S was given a sheet of paper and requested to recall and draw as many of the designs as he could. In the *non-anxiety group* (NA) Ss took the Bender in the sequence outlined above without the prior experience of repeated test failures. As a check on the effectiveness of the induced anxiety condition, each S was asked to complete a self-rated anxiety scale dealing with the degree of tension he felt during the experiment.

Measures. Determination of size deviations from the standard designs followed Kitay's [6] method. For each S the algebraic mean noting direction and size of deviation was calculated for both initial and recall performance. The self-rating of anxiety was scored by assigning an appropriate ordinal number to the point at which the S marked the scale.

Results

Table 1 reports the means and difference between the means for both the initial and recall phases of the Bender-Gestalt data. Analysis of variance of deviations in the initial reproduction phase indicates a statistically significant main effect for the

Table 1. Means of Size Deviations in Inches During Initial and Recall Phases and the Difference Between the Means.

Groups	Initial	Recall	Difference
High Taylor IA	−0.80″	−1.23″	.43″*
Low Taylor IA	−0.53″	−1.01″	.48″*
High Taylor NA	−0.22″	−0.69″	.47″*
Low Taylor NA	−0.15″	−0.76″	.59″*

*Significant beyond .05 level of confidence.

experimentally induced anxiety variable. The means were further subjected to the Duncan Multiple Range Test [2] and the results show that the deviation of −.80″ obtained by the Taylor-High IA group is significantly greater than the deviation of both High (−.22″) and Low (−.15″) NA groups, but not different than that (−.53″) of the Taylor-Low IA group.

Size deviation scores for the recall drawings reveal the same results as those found with the initial reproductions. The deviation of −1.23″ obtained by the Taylor-High IA group is significantly greater than the deviation of both High (−0.69″) and Low (−0.75″) NA groups. A direct analysis of the question of whether the recall phase was more discriminating of anxiety than the initial reproduction phase showed, by analysis of variance, no significant differences in performance between the groups. It is of interest to note, however, that although these difference scores showed no discriminability among groups, the size difference between the initial reproduction and the recall means was significant for all groups and indicated an overall shift toward smaller drawings in the recall phase.

Conclusions

Our results indicate that subjects working under experimentally induced conditions of anxiety produce significantly smaller Bender reproductions than those performing in nonstress conditions. Although the association between anxiety and size

decrements was demonstrated more clearly with the induced rather than the personality variable, closer analysis of the results showed the Taylor-High scorers contributing more to this effect than those in the Taylor-Low group.

The failure to find significant differences between the reproduction and recall phases of the Bender rejects the hypothesis that the recall section is more sensitive to subject differences, at least as gauged by our measure of anxiety. The significant decrease in the size of drawings during the recall phase suggests perhaps that this feature of the test may elicit anxiety beyond that normally evoked during the initial reproduction task. This hypothesis would be consistent with other elements of our findings and obtains support from frequently verbalized comments to that effect by our subjects. If this impression is correct, care should be taken not to infer that the recall phase is more sensitive to inner dispositions of anxiety; more accurately, it is an anxiety producing experience in itself.

References

1. Bender, L. A. A Visual Motor Gestalt Test and Its Clinical Use. *American Orthopsychiatric Association, Res. Monog.*, 1938, No. 3.
2. Duncan, D. B. Multiple Range and Multiple F Tests. *Biometrics*, 1955, Vol. II, No. 1.
3. Gobetz, W. A Quantification of the Bender-Gestalt Test on Normal and Neurotic Adults. *Psychol. Monog.*, 1953, *67*, No. 6 (Whole No. 356).
4. Hutt, M. L. Revised Bender Visual-Motor Gestalt Test. In Weider, A. (Ed.). *Contributions Toward Medical Psychology*. New York: Ronald Press Co., 1953, Vol. 11, 660-687.
5. Kitay, J. I. The Bender-Gestalt Test as a Projective Technique. *J. clin. Psychol.*, 1950, *6*, 170-74.
6. Taylor, J. A. A Personality Scale of Manifest Anxiety. *J. abnorm. soc. Psychol.*, 1953, *48*, 285-290.
7. Tolor, A. A Comparison of the Bender-Gestalt Test and the Digit-Span Test as Measures of Recall. *J. consult. Psychol.*, 1956, *20*, 305-309.

CHILDREN'S SPONTANEOUS OBJECT ELABORATIONS ON A VISUAL-MOTOR TEST

G. H. ZUK[1]

St. Christopher's Hospital for Children, Philadelphia

The object of this paper is to call attention to a regressive or primitive feature, apparently unrelated to test stimuli, that may occur in performance of a visual-motor test. In so far as it bears on the problem of diagnosis, it should be helpful to make some distinction between this feature, called for present purposes "spontaneous object elaboration", and those that are more obviously related to test stimuli. Bender (1) has offered numerous examples of primitivizations occurring on her visual-motor gestalt test, noting the tendency of sensorimotor patterns to revert back to more primitive patterns and express movement of some kind. She asserts that at all levels all of the original sensorimotor principles are in evidence. The writer (4) recently offered evidence supporting this view: children were found to employ a single drawing pattern in reproducing geometric figures requiring different levels of sensorimotor and conceptual ability for successful performance.

Procedure

In a previous study the writer (3) described the group administration to adolescent and adult retarded patients of a specially designed visual-motor test. This test was also administered to a somewhat younger group of retarded subjects whose mental age was believed to be below that required for successful performance. These children, 58 in number, attended the "primary grade" in the school at Pacific State

[1]Psychologist in the Mental Retardation Clinic, which is partly supported by grants from the Children's Bureau, Department of Health, Education and Welfare, and the Pennsylvania State Department of Health, Division of Maternal and Child Health.

Hospital,[2] an institution for mental defectives. In the classroom situation they were asked by the writer to copy one at a time 10 geometric designs. The writer stood at the front of the class, informed the students that he would expose one design at a time briefly (printed on a 5 x 8 inch white card), then retract the design and have them try to reproduce it.

Four conditions of administration consuming approximately one-half an hour were involved. In the first, each design was exposed for about 10 seconds and then retracted. The children attempted to copy the design on one side of a sheet of yellow paper (8½ x 11 inches) supplied them. When the 10 designs were completed, the papers were picked up and a second sheet supplied. On these the children again attempted to copy the designs, this time having seen each for approximately 5 seconds. On a third sheet of paper, the children copied the figures, having seen each this time for approximately 2 seconds. On a fourth sheet of paper, the children copied as many of the designs as they could recall from memory.

The children were given as much time to draw as seemed to be needed. The writer presented a new design after having assured himself that all children had stopped applying pencil to paper and were looking up. In the recall condition, the children handed in their papers when they agreed they could recall no more. The factor of not having sufficient time to copy was thus not felt to be influential in determining distortions or elaborations of the designs. Also, any child was permitted to sit close to the examiner who complained of not being able to see.

In examining the 58 test records it was found that all contained examples of regressive or primitive features as described by Bender and others. Surprisingly, however, another type of feature was encountered that appeared to have little actual relation to the test stimuli. Many of the children were found to have drawn human figures or human details (hands, heads), letters and numbers on the sheets of paper. Of the 58 records, 18 contained examples of this feature. Of these 18, information as to MA and CA of the child was available for 13. Of the remaining 40, information as to MA and CA of the child was available for 24. Table 1 presents this information for the 37 children composing both groups. It is believed that the 21 children for whom information was not available did not significantly affect the MA and CA composition of the groups reported in Table 1, inasmuch as all 58 children were in the "primary grade" at the School and all of them were in the MA and CA range reported for the group of 37.

Results

From Table 1 it seems clear that neither MA or CA were the factors separating the 13 children who drew spontaneous object elaborations from the 24 who did not.

Table 1. Some Descriptive Information on Children Who Did and Did Not Elaborate Objects on the Visual-Motor Test

Groups	Sex		Chronological age		Mental age	
	Male	Female	Mean	Range	Mean	Range
Children who elaborated objects (N = 13)	6	7	11.23	9-7 to 13-10	5.34	4-6 to 7-0
Children who did not elaborate objects (N = 24)	14	10	11.24	8-2 to 14-7	5.25	3-7 to 7-3

[2]The writer wishes to thank the following persons at Pacific State Hospital in Pomona, California, for their help in the course of the study: George Tarjan, medical director; Anna Shotwell, chief psychologist; and Mayme Dinsmore, principal of the school.

It seems clear, however, that elaborations of the type described were relatively specific to the children attending the "primary grade" at the school; that is, children in the MA range from 4 through 7 years. None of the children or adults in the higher-level classes drew this type of elaboration.

As previously noted, 18 of the 58 records examined contained examples of spontaneous object elaboration. Of these 18, 11 showed one elaboration in but one of the four conditions. Of these 11, 1 showed the feature in the 10 second condition, 0 in the 5 second condition, 3 in the 2 second condition and 7 in the total recall condition. Of 7 who showed elaborations in more than one of the four conditions, 1 showed it in all four, 1 in the first two, 4 in the last three and 1 in the last two.

The 18 records thus contained a total of 31 examples of the feature. Table 2 shows the distribution of these 31 features in the various conditions by type. In the last column of Table 2, it can be seen that 10 elaborations were of human figures or human details, 14 were of letters, numbers or words, 3 were of a combination of the first two categories, and 4 were of scenes or objects not included in the first two categories. In the last row of Table 2 can be seen frequencies of the 31 features in various conditions. Observed in the 10 second condition were 3; in the 5 second condition, 6; in the 2 second condition, 9; and in the total recall condition, 13.

TABLE 2. CATEGORIES OF ELABORATIONS MADE BY 18 CHILDREN AND THEIR FREQUENCIES IN VARIOUS CONDITIONS

	Conditions				
Category	10 sec.	5 sec.	2 sec.	Total Recall	Sum
Human figure or detail	2	1	3	4	10
Letters, numbers or words	1	3	5	5	14
Combination of first two categories	0	1	1	1	3
Scenes or objects other than in first two categories	0	1	0	3	4
Sum	3	6	9	13	31

Of the 18 records, 2 showed a greater frequency of elaboration in the first two conditions (10 second and 5 second) than in the last two (2 second and total recall); 2 showed an equal frequency in the first two and last two conditions; but 14 showed a greater frequency in the last two than first two conditions. A test[3] of the significance of this distribution in the various conditions resulted in a chi square of 16.0 which, with 2 degrees of freedom, is likely to occur less than one in a thousand times by chance. Object elaborations were thus shown to have occurred more frequently in the last two than first two conditions, and especially in the total recall condition.

DISCUSSION

Observations of spontaneous object elaborations in children taking a visual-motor test have been presented. Under the test conditions specified, retarded children drew recognizable objects (such as human figures, numbers and letters) which appeared to be functionally unrelated to the test stimuli. Past descriptions of such types of primitivizations have characteristically emphasized their relatedness to test stimuli.

[3] A chi square one-sample test was employed following the model supplied by Siegel (3).

Spontaneous object elaboration seems relatively unique among the various kinds of primitivizations reported to occur on visual-motor tests. Bender has demonstrated the value of these features for the diagnosis of mental retardation, brain injury and psychosis in children and adults. The feature described here seems primarily to have been a function of mental level and therefore useful in the assessment of mental functioning and also as a contrast to features said to be more typical of brain injury or psychosis.

Summary

Spontaneous elaboration of human figures and human details, letters, numbers and other recognizable objects was observed in 18 of 58 records of retarded children who were instructed to reproduce geometric designs exposed to them briefly. These elaborations occurred more frequently when children were asked to reproduce the designs from total recall than when they had an opportunity to see the designs just prior to copying them.

The children producing the elaborations were in the MA range from 4 through 7 years. The explanation is offered that they represented a concrete substitutive behavior occurring in the absence of necessary sensorimotor and conceptual skill. The diagnostic significance of the feature was noted, especially in contrast to features outlined by Bender as more typical of brain injury and psychosis. The unique aspect of the feature—its apparent unrelatedness to test stimuli—was pointed out. The use of repeated test conditions, and especially a total recall condition, was felt to be crucial to production of the feature.

References

1. Bender, Lauretta. *A visual motor gestalt test and its clinical use.* New York: Amer. Orthopsychiat. Ass., 1938.
2. Siegel, S. *Nonparametric statistics.* New York: McGraw-Hill, 1956.
3. Zuk, G. H. The mind as an optical system: a study of perceptual reproduction in the normal and mentally retarded. *J. genet. Psychol.*, 1959, *94*, 113-130.
4. Zuk, G. H. Size: its significance in the copied drawings of children. *J. clin. Psychol.*, 1960, *16* 38-42.

A FACTOR ANALYTIC STUDY OF BARRON'S EGO STRENGTH SCALE[1]

EVELYN CRUMPTON — *Veterans Administration Neuropsychiatric Hospital and University of California Medical Center, Los Angeles*

JOEL M. CANTOR — *Veterans Administration Neuropsychiatric Hospital and College of Medical Evangelists, Los Angeles*

CURT BATISTE — *University of California Medical Center, Los Angeles*

Since Barron's Ego Strength (Es) scale[2] for the MMPI first appeared, an impressive array of validating evidence has been reported, both by Barron and by independent investigators, suggesting that ego strength is indeed what is measured by the Es scale. In addition to the original criterion variable of improvement with psychotherapy[2, 5], correlates of high Es scores are such variables as low MMPI

[1]From the Veterans Administration Neuropsychiatric Hospital, the College of Medical Evangelists, and the University of California Medical Center, Los Angeles. This study was supported in part by a Public Health Service fellowship grant and by research funds of the Departments of Psychiatry, University of California Medical Center and the College of Medical Evangelists. Statistical analysis of data was performed on Standards Western Automatic Computer, through the courtesy of the Office of Naval Research and the Institute for Numerical Analysis at the University of California, Los Angeles. The authors wish to express their grateful appreciation to Dr. Andrew L. Comrey, who made available his programs for SWAC; to Dr. John R. Peters and Dr. Harry E. Edwards, who furnished the data on medical students; and to Dr. Wilfred J. Dixon, who acted as statistical consultant. A shorter version of this paper was presented at the Western Psychological Association meeting, San Diego, 1959.

clinical symptom scales, intelligence, and lack of ethnic prejudice[3]; ratings of poise, drive, etc.[2]; Q-sort items descriptive of personality functioning[4]; independence of judgment in group situational test[1]; ability to orient oneself correctly to the vertical plane in darkness[13]; favorable aspects of functioning such as W and M on the Rorschach[11]; denial of need for help[6]; improvement in hospitalized patients[12].

One is struck by the almost complete lack of contradictory or inconsistent findings[10, 9]. On the other hand, data yielded in our own preliminary work with the scale suggested that further refinement was needed. In one small study, therapeutic improvement rated by psychiatrists in a group of 26 outpatient psychoneurotics was found to be inversely correlated ((rho = −.41, p = .05) with Es scores obtained prior to therapy, and not significantly related (rho = −.08) to changes in Es scores obtained before and after therapy. In another preliminary study, responses to 43 of the Barron items were determined from administration of the short form MMPI to 19 hospitalized psychotics before and after a course of shock treatment. Although pre-treatment Es score was not significantly related to subsequent clinical improvement (rho = .01), the changes in Es scores did conform to what would be expected, *i.e.*, improvement was reflected in higher Es scores at the end of treatment (rho = .60).

The Es scale thus appears to be a promising tool for research and clinical use, well worthy of further development. The purpose of the present study is to determine the scale's factorial structure so that the underlying construct may be better understood and the accuracy of its measurements be increased through scoring refinements.

Method

Subjects. Subjects consisted of 110 male veteran hospitalized psychiatric patients and 70 male medical students. The patient group consisted exclusively of open ward patients, predominantly schizophrenic in diagnostic category, with an admixture of depressives, character disorders, and severe chronic psychoneurotic disorders. The students came from a medical school with a religious emphasis. The total group ranged in age from 20 to 76, with a mean of 33. The patients were somewhat older than the students, the patient mean being 37 and the student mean 26. By virtue of the selection criterion, the group of medical students, of course, exceeded the group of patients in educational level, as well as in personal adjustment as evidenced by lack of hospitalization.

Procedure. The machine scoring form of the MMPI was administered to subjects in small groups. The test records were used to ascertain the responses of each subject to each of the items of the Es scale, as well as the total Es raw score.

The statistical procedures, described more fully in the Results section, are summarized as follows: The responses to the Es items were intercorrelated and factor analyzed. The extracted factors were then treated as sub-scales with each subject being given scores for each factor. The total raw scores and the individual factor scores were studied to see which, if any, differentiated students and patients. Ambiguous and non-contributory items, identified by inspection of the factor loadings and appearance only in non-differentiating factors, were eliminated from the scale. The two groups—students and patients—were then compared with respect to the new Es total scores yielded by the revised scoring system.

Results

Responses of the 180 subjects to the 62 items of the Es scale[2] and the criterion variable of student or patient were individually tabulated, and all possible item inter-

[2]Since the SWAC program could accommodate only 63 variables, six items (2, 36, 132, 187, 231, 236) were omitted, these items being selected at random with a double check being made to insure that the content of the eliminated item was represented in similar form in the remaining items. The 63 variables thus consisted of 62 Es items and the criterion variable of student or patient status.

correlations were determined. The matrix of phi coefficients is shown in Table A.[3] This correlational matrix was factor analyzed by the Thurstone centroid method and rotated analytically by the Kaiser Varimax method[7]. Fourteen factors were extracted, the last two having no centroid loading over .24. Table B shows the centroid factor loadings, Table C the rotated factor loadings. Some highly tentative identifications of the factors have been made by considering the possible common element or elements of items with factor loadings of .25 or above for a particular factor, elements not shared by items with near zero loadings. The adequacy of the naming varies considerably from factor to factor; these identifications must be considered as no more than speculative.

Factor scores were ascertained for each subject. The two subgroups, students and patients, were then compared with respect to each factor. Table 1 gives the means for the two groups and the results of the t test for the significance of differ-

TABLE 1. RESULTS OF THE COMPARISON OF STUDENTS AND PATIENTS ON FACTOR SCORES AND TOTAL Es SCORES

Factor	Patient Mean	Student Mean	t*	Cut-off Point for Classification as Student	% Correctly Classified Patient	% Correctly Classified Student
I	8.62	10.50	6.30	10 or more	56.4	91.4
II	4.14	1.07	13.80	2 or less	83.6	91.4
III	1.15	2.03	5.81	1 or more	66.4	100.0
IV	10.54	13.73	5.91	11 or more	45.4	95.7
V	3.21	4.73	7.76	4 or more	54.5	90.0
VI	1.57	1.36	1.04	—	—	—
VII	4.62	6.00	6.19	5 or more	40.0	88.6
VIII	2.98	0.39	8.39	2 or less	51.8	100.00
IX	2.90	3.54	4.20	4 or more	60.9	64.2
X	2.55	1.85	3.58	2 or less	50.9	71.4
XI	1.29	0.27	5.52	zero	58.2	75.7
XII	3.71	4.71	6.41	4 or more	37.3	98.6
XIII	3.35	4.34	5.51	4 or more	48.2	92.9
XIV	4.01	5.07	6.09	5 or more	61.8	77.1
Barron Total Scale	38.24	45.58	7.73	40 or more	54.5	97.2
Revised Total Score	33.66	42.86	9.43	40 or more	74.5	92.9

*With the exception of Factor VI, all t values are significant at beyond the .001 level.

ences between each pair of means. It will be noted that in only one factor (VI) was there no significant difference between the means of patients and students. Medical students had higher means on nine factors, psychiatric patients on five factors.

The discriminatory value of the individual factor scores was tested further. The most effective cut-off score was determined by inspection of the two distributions (students and patients) of scores for each factor. Table 1 indicates for each factor the cut-off score and the percentages of patients and students correctly classified by use of the cut-off score. For each factor, the percentage correctly classified by chance is 61 for patients, 39 for students. The results given in Table 1 make it clear that students are easily identified by use of any of the factors (except VI); however, identification of patients as patients is by no means that clearcut. It was thought that the patterning of the factor scores might help in discriminating patients with

[3]Tables A, B, and C have been deposited with the American Documentation Institute. Order Document No. 6147 from the ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C., remitting in advance $1.25 for 35 mm Microfilm or $1.25 for 6 x 8 photocopies. Make check payable to Chief, Photoduplication Service, Library of Congress.

high ego strength scores from the student group. When the data were inspected, however, this approach appeared no more promising than the use of the individual factor scores.

Factor I. Factor I appears to be a general factor relating to absence of symptomatology. Eleven items had loadings over .25, seven of these have loadings over .40. The mean for the student group was higher than the mean for the patient group on this factor, as would be expected. Three of these items appear in Barron's "sense of reality" category, four in "phobias, infantile anxieties".

Item	Loading	Description
541	58	My skin seems to be unusually sensitive to touch. (F)
48	52	When I am with people, I am bothered by hearing very queer things. (F)
510	49	Dirt frightens or disgusts me. (F)
494	48	I am afraid of finding myself in a closet or small closed space. (F)
341	43	At times I hear so well it bothers me. (F)
420	42	I have had some very unusual religious experience. (F)
525	42	I am made nervous by certain animals. (F)
43	29	My sleep is fitful and disturbed. (F)
555	27	I sometimes feel that I am about to go to pieces. (F)
22	26	At times I have fits of laughing and crying that I cannot control. (F)
33	26	I have had very peculiar and strange experiences. (F)

Factor II. The second factor has to do primarily with religious attitudes. This is one of the few factors on which the item, patient or student, had a high loading. One might hypothesize that the difference between patients and students on these items is that the patients with good ego strength express reasonable religious feelings without fanaticism, whereas students at a presumably much higher level of ego strength are able to tolerate some rather dogmatic attitudes. In connection with this factor, it should be remembered that the students were enrolled in a medical school with a religious emphasis, making this student group not representative in terms of religious attitudes. The scores for factor II were calculated with the item, patient or student, omitted. The mean for the patients was higher than the mean for the students on the factor. Four items appear in Barron's category, "attitudes toward religion". The items are:

Item	Loading	Description
95	72	I go to church almost every week. (T)
488	70	I pray several times every week. (F)
	−67	Patient or student.
58	61	Everything is turning out just like the prophets of the Bible said it would. (F)
378	55	I do not like to see women smoke. (F)
483	47	Christ performed miracles such as changing water into wine. (F)
554	32	If I were an artist I would like to draw children. (F)

Factor III. The third factor again included the criterion item, student or patient. The content of the remaining four items in this factor would seem to have a somewhat psychopathic flavor, a rebelliousness that might conceivably be related to superior ego strength in a high ego strength group and to ego defect in a weak ego group. With the patient-student item omitted, the mean for the student group is higher than the mean for the patient group. Items with factor loadings over .25 were:

Item	Loading	Description
410	51	I would certainly enjoy beating a crook at his own game. (T)
109	49	Some people are so bossy that I feel like doing the opposite of what they request, even though I know they are right. (T)
355	28	Sometimes I enjoy hurting persons I love. (T)
251	−27	I have had blank spells in which my activities were interrupted and I did not know what was going on around me. (F)
	25	Patient or student.

Factor IV. Factor IV has sixteen items with loadings over .25, the items appearing to be somewhat similar to the items in Factor I. The content of Factor I

items refers to a fairly wide variety of symptoms. Factor IV items have to do primarily with anxious, ruminative, worrying, obsessive, distractable feelings or behaviors, the absence of such making for a high score on the factor. As would be expected, students had a higher mean factor score than patients. Seven of these sixteen items appear in Barron's category, "psychasthenia and seclusiveness", four in "personal adequacy and ability to cope". Items with loadings of over .25 were:

Item	Loading	Description
389	63	My plans frequently seem so full of difficulties that I have had to give them up. (F)
94	55	I do many things which I regret afterwards. (I regret things more or more often than others seem to.) (F)
344	54	Often I cross the street in order not to meet someone I see. (F)
359	52	Sometimes some unimportant thought will run through my mind and bother me for days. (F)
100	52	I have met problems so full of possibilities that I have been unable to make up my mind about them. (F)
217	51	Frequently I find myself worrying about something. (F)
555	48	I sometimes feel that I am about to go to pieces. (F)
32	47	I find it very hard to keep my mind on a task or job. (F)
234	−46	I get mad easily and then get over it soon. (T)
251	43	I have had blank spells in which my activities were interrupted and I did not know what was going on around me. (F)
349	42	I have strange and peculiar thoughts. (F)
241	40	I dream frequently about things that are best kept to myself. (F)
384	37	I feel unable to tell anyone all about myself. (F)
244	35	My way of doing things is apt to be misunderstood by others. (F)
355	−32	Sometimes I enjoy hurting persons I love. (T)
153	27	During the past few years I have been well most of the time. (T)

Factor V. The identification of Factor V is quite uncertain. Six items had loadings over .25, and there is no obvious common thread running through the content of these items. One attractive speculation is that the items represent a denial of weakness or of a "sissified" attitude in the face of stress. Disaffirming the hypothesis, however, is the near zero loading of item 261. Items were:

Item	Loading	Description
174	47	I have never had a fainting spell. (T)
251	41	I have had blank spells in which my activities were interrupted and I did not know what was going on around me. (F)
554	35	If I were an artist I would like to draw children. (F)
378	−32	I do not like to see women smoke. (F)
548	−32	I never attend a sexy show if I can avoid it. (F)
62	28	Parts of my body often have feelings like burning, tingling, crawling, or like "going to sleep". (F)

The first five factors are probably the most meaningful of the group. However, the additional factors extracted should not be ignored.

Factor VI. Factor VI consists of four items with seemingly dissimilar content. This was the only one of 14 factors of which there was not a significant difference in the mean of the two groups. One attitude that all four items perhaps have in common is a lackadaisical, take-it-easy, non-assertive approach. Items for the loadings of over .25 are as follows:

Item	Loading	Description
380	−53	When someone says silly or ignorant things about something I know about, I try to set him right. (T)
32	−28	I find it very hard to keep my mind on a task or job. (F)
561	27	I very much like horseback riding. (F)
384	−26	I feel very unable to tell anyone all about myself. (F)

Factor VII. For this factor, seven items had loadings of .25 or over. The first item, with a factor loading of .60, contributes most to the factor. One would certainly expect medical students to like science more than a less occupationally homogeneous group such as the patients. On the face of it, one might expect that this factor is a

result of the particular sample, rather than being a real quality of ego strength. The fact that the patient-student item does not have a loading of at least .25 for this factor, however, argues for the opposite hypothesis that factor VI does reflect ego strength to some extent. The content of each of the remaining items has the flavor of righteousness, assertive virtue, strong convictions. Students exceeded patients in mean factor score. The items are:

Item	Loading	Description
221	60	I like science. (T)
82	38	I am easily downed in an argument. (F)
95	−31	I go to church almost every week. (T)
94	28	I do many things which I regret afterwards. (I regret things more or more often than others seem to.) (F)
525	28	I am made nervous by certain animals. (F)
355	−27	Sometimes I enjoy hurting persons I love. (T)
384	25	I feel unable to tell anyone all about myself. (F)

Factor VIII. Factor VIII items have to do with weakness, primarily a physical lassitude, somatizing and tension. The mean of patients is higher than the mean of students for this factor. Four of these items appear in Barron's category "physical functioning and physiological stability". Items with loadings of .25 or over are as follows:

Item	Loading	Description
189	−70	I feel weak all over much of the time. (F)
544	−60	I feel tired a good deal of the time. (F)
153	−47	During the past few years I have been well most of the time. (T)
62	−43	Parts of my body often have feelings like burning, tingling, crawling, or like "going to sleep". (F)
51	−34	I am in just as good physical health as most of my friends. (T)
43	−31	My sleep is fitful and disturbed. (F)
389	−28	My plans frequently seem so full of difficulties that I have had to give them up. (F)
555	−28	I sometimes feel that I am about to go to pieces. (F)
192	−25	I have had no difficulty in keeping my balance in walking. (T)

Factor IX. Four items had loadings over .25 for this factor. The item making the greatest contribution to the factor is concerned with the subjects' having had a reasonably adequate home life, and this is implied to a certain extent by the other three items. Students had a higher mean factor score than patients. The four items are:

Item	Loading	Description
515	54	In my home we have always had the ordinary necessities (such as enough food, clothing, etc.). (T)
253	42	I can be friendly with people who do things which I consider wrong. (T)
559	37	I have often been frightened in the middle of the night. (F)
421	−27	One or more members of my family is very nervous. (T)

Factor X. The meaning of Factor X is certainly not obvious from the content of the five items with loadings of .25 or over. Patients exceed students in factor scores on these items. One might speculate that a person who answered each of these items in the indicated direction would be quite passive and immature, lacking in aggressiveness. (In this regard, Barron found "effeminate" checked for low Es scores.) The items are:

Item	Loading	Description
261	−61	If I were an artist I would like to draw flowers. (F)
554	−38	If I were an artist I would like to draw children. (F)
208	−36	I like to flirt. (T)
367	−28	I am not afraid of fire. (T)
458	−25	The man who had most to do with me when I was a child (such as my father, step-father, etc.) was very strict with me. (T)

Factor XI. Factor XI has five items with loadings over .25, most of them suggestive of an active and disturbing fantasy life. The mean of the patient group is higher than the mean of the student group in this factor. The items are:

Item	Loading	Description
43	−54	My sleep is fitful and disturbed. (F)
349	−31	I have strange and peculiar thoughts. (F)
241	−40	I dream frequently about things that are best kept to myself. (F)
51	−32	I am in just as good physical health as most of my friends. (T)
33	−28	I have had very peculiar and strange experiences. (F)

Factor XII. There are five items with loadings over .25 on factor XII. What these items have in common is hard to determine. The group of medical students has a higher mean than the group of patients. The items are:

Item	Loading	Description
209	52	I believe my sins are unpardonable. (F)
34	39	I have a cough most of the time. (F)
244	31	My way of doing things is apt to be misunderstood by others. (F)
192	27	I have had no difficulty in keeping my balance in walking. (T)
62	26	Parts of my body often have feelings like burning, tingling, crawling, or like "going to sleep". (F)

Factor XIII. No attempt has been made to interpret this factor. Students are higher in mean factor score than patients. The five items with factor loadings over .25 are:

Item	Loading	Description
140	40	I like to cook. (F)
33	34	I have had very peculiar and strange experiences. (F)
341	28	At times I hear so well it bothers me. (F)
62	27	Parts of my body often have feelings like burning, tingling, crawling, or like "going to sleep". (F)
525	27	I am made nervous by certain animals. (F)

Factor XIV. The six items with loadings of .25 or over for Factor XIV are somewhat varied in content, the two most heavily loaded items having to do with heterosexuality. The mean of the medical student group is higher than the mean of the patient group. The items are:

Item	Loading	Description
430	51	I am attracted by members of the opposite sex. (T)
208	37	I like to flirt. (T)
14	33	I have diarrhea once a month or more. (F)
270	29	When I leave home I do not worry about whether the door is locked and the windows closed. (T)
192	25	I have had no difficulty in keeping my balance in walking. (T)
510	25	Dirt frightens or disgusts me. (F)

Revision of Scoring. The naming of the factors is considered of secondary importance, primary interest being in the usefulness of the factor information in refining the scoring of the scale. As would be expected, the group of students had a higher mean ego strength score for the 62 item scale than did the group of patients (see Table 1). This difference was significant at beyond the .001 level. The cut-off point giving the highest number of subjects correctly identified was 40 or more for classification as a student, 39 or less as a patient. With this cut-off point score used for prediction, 54.5% of the patients were correctly identified, 97.2% of the students. This cut-off point obviously yields a percentage of correct classification that is quite acceptable for groups. For the identification of individuals, particularly in the patient group, discrimination could be improved.

The results of the factor analysis suggested that the scoring of some items decreased the utility of the scale for this sample. The scoring system was therefore revised on the basis of information yielded by the factor analysis. The scoring of three items was reversed (27, 46, 60), and 12 items were omitted (10, 16, 20, 30, 37, 40, 44, 49, 50, 51, 54, 63). Omitted items were those that appeared only in factors where there appeared to be a difference in interpretation according to whether the subject was a patient or a student, and a few items scored false for one Es factor and true for another. Items were scored in reverse for items whose opposite scoring was associated with an Es factor.

The new scale thus consisted of 50 items, rather than 62. The difference between the means of the revised scores was also high enough for significance at beyond the .001 level (see Table 1). Thus the revised scoring does at least as well as the Barron scoring in separating the means of the two distributions; it was not possible to test statistically whether the scoring revision more effectively separated the means. As a matter of interest, the *t* tests reported in Table 1 were repeated using revised factor scores. The *t* values were apparently quite similar for the two methods of scoring.

Again, the cut-off point most effectively predicting classification of the subjects was 40 or above for students, 39 or below for patients. This cut-off score resulted in 74.5% of the patients correctly identified, and 92.9% of the students correctly classified. Compared with the original scoring, the revised scoring system was apparently somewhat more effective (74.5% to 54.5%) in discriminating patients from students and only slightly less effective (92.9% to 97.2%) in discriminating students from patients.[4]

Discussion

The results of the factor analysis and the subsequent treatment of the data suggest strongly that the Barron ego strength scale is misnamed. Inspection of the content reveals that the scale takes a negative approach to the measurement of ego strength, *i.e.*, the absence of indications of ego weakness is interpreted as ego strength.

Few of the medical students admit to signs of ego weakness such as are contained in the scale, e.g., symptoms, dogmatic attitudes, rigidities, anxieties; and almost all of the students achieved high ego strength scores. Many of the patients admit to many of these ego defects. But too many of the patients, like the medical students, do not admit to these ego weaknesses. Perhaps their psychopathology or ego defects are revealed by other indications not included in the scale.

The hypothesis that many of these hospitalized patients are in fact equal in basic ego strength to the medical students who are able to function in a non-protected environment outside of a hospital, does not seem to be a very tenable hypothesis. Certainly there must be positive qualities of ego strength that would separate more clearly the effective students from the non-effective patients. Clinicians are ordinarily interested in assessing the ego strength of the not-so-effective individual, not of the person who is demonstrating an adequate level of ego strength in his every day behavior. That is, from a practical standpoint, the percentage of correct classification for the patient group should be high; whereas it is not so important to have a high percentage of correct classification of the students.

In other studies, there has been enough construct validation of Barron's scale to suggest that the scale does have something to do with ego strength. The results of the present study, however, would seem to indicate that "ego weakness" would be a better term of what the scale is measuring than the term "ego strength". This is not merely a quibbling over names. The construct of ego strength should imply more than merely the absence of specific weaknesses. Perhaps the really strong ego can tolerate quite a few weaknesses in its otherwise very adequate structure. In this respect it is of interest to note that Es was loaded −.73 for the factor identified as "Ego Weakness vs. Ego Strength" in the factor analysis of the MMPI scales reported by Kassebaum *et al.* (8)

Summary

1. The Barron Ego Strength scale of the MMPI was factor analyzed by the Thurstone centroid method and rotated analytically by the Kaiser Varimax method. Subjects were 110 male veteran hospitalized psychiatric patients and 70 male medi-

[4]Since the cut-off point was chosen for maximal discrimination, chi square values take advantage of chance fluctuations and are thus inflated to an unknown degree. However, it may be of interest to note that the chi square value associated with the contingency table based on Barron scoring is 50.6, on the revised scoring 77.8. Both values are obviously very significant statistically.

cal students. Fourteen factors were extracted, 13 of which significantly discriminated the means of the two groups. Predictive efficiency of cut-off points on these factors was quite good for students and no more than fair for patients.

2. The results of the factor analysis were used to identify ambiguous, misleading, or non-contributory items, for the purpose of refining the scoring system. The results suggested omission of 12 items and reversal of scoring of three items.

3. Results support the contentions of Barron and others that the Es scale is a valuable tool of personality evaluation. Because the scale efficiently associates good ego strength with the effective functioning of students, and less efficiently associates poor ego strength with the non-effective functioning of psychiatric patients, it is proposed that what is being measured is the absence of specific ego weaknesses, and not the presence of ego strength.

References

1. Asch, S. Effects of group pressure upon the modification and distortion of judgments. Swarthmore College: Progress Report on Office of Naval Research Project, Task Order 70, MR-38003, 1950.
2. Barron, F. An ego strength scale which predicts response to psychotherapy. *J. consult. Psychol.*, 1953, *17*, 327-333.
3. Barron, F. Some test correlates of response to psychotherapy. *J. consult. Psychol.*, 1953, *17*, 235-241.
4. Barron, F. Ego strength and the management of aggression. Berkeley: Univ. Calif. Institute of Personality, Assessment and Research, 1953c.
5. Barron, F. and Leary, T. Changes in psycho-neurotic patients with and without psychotherapy. *J. consult. Psychol.*, 1955, *19*, 239-245.
6. Dymond, R. F. Adjustment changes in the absence of psychotherapy. *J. consult. Psychol.*, 1955, *19*, 103-107.
7. Kaiser, H. F. An analytic rotational criterion for factor analysis. Presented at the Amer. Psychol. Assn., Sept., 1955.
8. Kassebaum, G. C., Couch, A. S., and Slater, P. E. The factorial dimensions of the *MMPI*. *J. consult. Psychol.*, 1959, *23*, 226-236.
9. Pascal, G. R. and Suttell, B. J. *Bender-Gestalt test: Its qualification and validity for adults.* New York: Grune & Stratton, 1951.
10. Tamkin, A. S. Evaluation of the construct validity of Barron's Ego Strength scale. *J. clin. Psychol.*, 1957, *13*, 156-158.
11. Williams, H. L. and Lawrence, J. F. Comparison of the Rorschach and MMPI by means of factor analysis. *J. consult. Psychol.*, 1954, *18*, 193-197.
12. Wirt, R. D. Further validation of the ego strength scale of Barron. *J. consult. Psychol.*, 1955, *19*, 444.
13. Witkins, H. A. *Personality through perception.* New York: Harpers, 1953.

THE REPRESENTATION OF MMPI SCALES BY MMPI FACTOR SCALES

EDWARD F. GOCKA AND HAYDEN L. MEES[1]

Veterans Administration Hospital, American Lake, Washington[2]

Problem

In 1959, Mees[14] factored two intercorrelation matrices representing a large number of non-overlapping items from all of the MMPI clinical scales and obtained several factors. The first three of these factors (M_1, M_2, and M_3) were of sufficient strength and stability to justify the construction of relatively pure factor scales from the heavily loaded items. All items included in these scales had similar factor loadings both for a group of 300 heterogeneous university counseling center clients and hospitalized neuro-psychiatric patients of mixed diagnoses. The Ss were equally

[1]Now at Western State Hospital, Fort Steilacoom, Washington.

[2]The authors wish to acknowledge the assistance of Andrew Chambers and Hildegund Holloway in computational analysis and organization; and John B. Marks for making available data on hospitalized patients and for his constructive review of the manuscript.

divided in each group by sex but were not systematically balanced on other demographic characteristics. This method of obtaining scale items stands in contrast to the more common method of factoring a particular MMPI scale, as well as being somewhat unique in selecting items common to two divergent groups. The question remained as to how well these factor scales could now predict the usual clinical scales for some specific group of Ss. Furthermore, the relationship of these factor scales to other selected non-clinical, factorial and non-factorial, scales was of interest from the point of view of scale identification.

Mees tentatively identified scale M_1 as a "sickness" or "general complaint" factor which seemed heavily loaded by the psychotic dimension, scale M_2 was tentatively identified as belonging to an "expressive-repressive" or "introversive-extraversive" dimension, while scale M_3 was said to consist of a pool of items indicating "femininity" or "tenderminded endorsements". It was almost immediately obvious that scale M_1 would be related to the "social desirability" factor that appears so strongly in the MMPI items, but scales M_2 and M_3 were not so clearly defined. Hence, a secondary purpose of this paper centered around the attempt to clarify these two scales.

Procedure

A multiple-R analysis was chosen to determine how well the three factor scales represented other MMPI scales. Multiple correlations of the M_1, M_2, and M_3 factor scales were obtained with all of the standard clinical and validity scales as well as the clinical-subtle subscales. In the interest of further clarification a selection of special MMPI scales was made by simply taking a fairly recent published list [12] of "new" MMPI scales; the Taylor manifest anxiety scale [18] was added to this list because of its frequency in the literature. The Welsh A and R factor scales [19] were considered in a separate section to distinguish them from the non-pure-factor, or rather multiple factor scales; along with these two was included the Edwards' social desirability scale [5] which has often been referred to as a measure of the most prominent factor in the MMPI. Four additional scales were included in yet another section in the attempt to give the M_2 scale a more concrete definition; however, the authors were unable to find any scales relating strongly to factor scale M_3.

The sample of Ss consisted of 84 hospitalized male patients who had a former diagnosis of schizophrenia but were now at the point of being discharged. The MMPI, as well as other tests, were given to this group as a standard discharge battery. The average S's age was 38 with the standard deviation of the group being 10.

Results

The results of this analysis are given in Table 1. The multiple correlation Rc has been adjusted for chance errors due to the number of cases and the number of predictor variables, using Snedecor's correction.

Standard symbolic nomenclature were used for identifying the MMPI clinical and validity scales as listed in Table 1. The subtle scales were given an "S" after their standard designations. The selected special MMPI scales are identified as follows: Manifest anxiety scale (Mas) [18]; Dependency (Dy) [15]; Caudality (Ca) [20]; Prejudice scale (Pr) [10]; Ego strength (Es) [1]; Dominance (Do) [8]; Role playing (Rp) [13]; Responsibility (Re) [9]; Social status scale (St) [7]; Control (Cn) [4]; and Low back pain (Lb) [11].

The factor scales in Table 1 consist of the Welsh anxiety scale (A) [19]; Edwards' 39 item social desirability scale (SD) [5]; and the Welsh repression scale (R) [19]. The M_2 identifier scales are listed in Table 1 as follows: the 47 MMPI items of the manifest hostility scale (MHS) [17]; Ego control scale number 3 (EC_3) [2]; WR-X—the scale of 25 "X" type items from the Welsh R scale; and Internalization-Externalization scale number 3 (IE_3) [16].

The correlations, multiple correlations and standard partial regression coefficients (β-weights) express the pattern of relationships found. The special non-

TABLE 1. PREDICTION OF MMPI SCALES FROM THE M_1, M_2, and M_3 SCALES

MMPI Scales	Product Moment Correlations with			Standard Partial Regression Coefficients			Multiple
	M_1	M_2	M_3	M_1	M_2	M_3	Rc
Standard Test							
L	−.35	.10	.23	−.30	.06	.16	.34
F	.41	−.20	−.15	.37	−.14	−.07	.40
K	−.79	.30	.24	−.74	.17	.07	.80
Hs	.66	.03	−.17	.70	.16	.02	.67
D	.41	.40	−.37	.44	.46	−.22	.65
Hy	−.10	.43	.11	.02	.45	.16	.42
Pd	.45	.03	−.15	.47	.11	−.02	.43
Mf	.16	.35	.11	.29	.43	.22	.44
Pa	.36	.04	−.08	.39	.11	.03	.33
Pt	.90	−.08	−.25	.91	.08	−.02	.90
Sc	.79	−.16	−.23	.78	−.02	−.04	.78
Ma	.42	−.55	.06	.36	−.48	.10	.63
Si	.53	.32	−.40	.55	.40	−.23	.70
Clinical-Subtle							
HyS	−.74	.29	.39	−.65	.20	.25	.79
DS	−.51	.35	−.04	−.50	.25	−.14	.57
MaS	.03	−.48	.29	.01	−.46	.25	.52
PdS	−.32	.14	.16	−.28	.10	.10	.29
PaS	−.14	.45	.05	−.04	.45	.08	.43
Special							
Mas	.89	−.05	−.28	.90	.11	−.05	.90
Dy	.88	.00	−.16	.94	.18	.09	.90
Ca	.83	.11	−.32	.86	.26	−.08	.87
Pr	.78	−.39	−.22	.72	−.27	−.07	.82
Es	−.74	.01	.13	−.79	−.14	−.08	.75
Do	−.68	.15	.35	−.62	.06	.20	.69
Rp	−.68	.06	.30	−.66	−.04	.13	.68
Re	−.63	.33	.24	−.56	.24	.13	.67
St	−.63	−.03	.36	−.60	−.12	.20	.66
Cn	.45	−.28	−.29	.35	−.24	−.23	.51
Lb	−.38	.20	−.15	−.42	.10	−.25	.43
Factor							
Welsh A	.87	−.11	−.23	.88	.05	−.01	.87
SD	−.81	.05	.23	−.83	−.10	.02	.81
Welsh R	−.25	.59	−.23	−.21	.53	−.23	.64
M_3 Identifiers							
MHS	.64	−.39	−.18	.57	−.30	−.07	.69
EC_3	−.22	.61	−.14	−.15	.57	−.12	.62
WR-X	−.05	.54	−.31	−.02	.51	−.27	.58
IE_3	−.35	.41	.09	−.27	.37	.06	.48

clinical scales obtained higher multiple R's on the average than did the clinical scales. The β-weights reveal that most of the variance of the non-clinical scales is obtained from the relationship with factor scale M_1. For the clinical and validity scales, K, Hs, Pt, Sc, and HyS show the strongest relationship with M_1. Factor scale M_2 does not have the large covariation with other scales as does M_1; M_2 relates in a moderate way to D, Hy, Mf, Ma, Si, MaS and Pa of the clinical scales. Factor scale M_3 relates only at low levels to D, Mf, Si, HyS, and MaS of the clinical scales.

Although M_1 is well represented in the special non-clinical scales, notable exceptions are the Cn, Lb, R, EC_3, WR-X, and IE_3 scales. Factor scale M_2 has a moderate relationship to the non-clinical scale R, and the M_3 identifiers, while factor scale M_3 has only small relationships with the non-clinical scales Lb, Cn, St, Do, R, and WR-X. For all of the scales the overall sizes of the β-weights seem to reflect the hierarchy of the original factors obtained by Mees.

Discussion

From this analysis we find that a good number of the selected non-clinical scales relate strongly to the M_1 factor scale, but only certain of the clinical scales relate strongly to this scale. In a previous factor analytic study[6] other scale relationships with the M_1 scale produced a second-order factor named General Complaint *vs.* Dynamic Integration. Highly implicit in this second-order factor is the tendency to respond in a socially desirable fashion; in fact the correlations of the clinical scales in Table 1, with M_1, show good correspondence to the relationship reported by Edwards[5] between the Social Desirability scale and the MMPI clinical scales (with signs reversed). Hence, the tendency to answer in a socially desirable fashion or a socially undesirable fashion is reflected in these correlations with M_1, and is revealed by the large negative correlation with the SD scale. Most appropriately, this tendency reflects the process of Dynamic Integration[3] at one pole with lack of Integration at the other pole which could be labelled "General Complaint." Of all the scales only Hy, MaS, PaS, EC_3, and WR-X have β-weights on M_1 smaller than .20. This indicates that the integration-complaint dimension enters into every clinical and validity scale of the MMPI by either relating to the total scale of items or to a subscale of items; *e.g.*, if M_1 does not relate to the total Hy scale it relates to the HyS subscale. The degree of relationship for the special non-clinical scales of this study, however, is consistently higher on the average than it is for the clinical and validity scales.

Factor scale M_2, on the other hand, is moderately represented in the clinical scales. Its representation in the non-clinical scales is highest on the repression factor scale (R) and on the contrived scales called M_2 identifiers which were included expressly to tap this dimension. These scales were found to group together on a second-order factor representing a Repressive-Expressive dimension[6]. Factor scale M_3 is represented in all scales only at a low level and must remain a relatively undefined scale at this point. In considering the total relations from all three M-scales, we might surmise that the developers of experimental scales based on some non-pathological attribute, usually fall heir to halo effects attendant upon the social desirability status of the items.

Perhaps for this reason, the combination of the three predictors accounts, on the average, for more of the selected non-clinical scale variance than it does of the clinical scale variance. When considering the nine standard clinical scales we find that less than .20 of their variance is accounted for in the case of the Hy, Pd, Mf, and Pa relationships with the predictors. On the other hand, for the non-clinical scales only Lb gives a multiple-R-squared of less than .20. From all indications it appears that at least one or two more factor scale predictors are necessary to raise the percentage of variance accounted for in the clinical scales to a reasonable level which would approximate the expected test-retest reliability variances. Until these additional scales are found there is little reason for substituting the proposed experimental scales in place of the usual clinical scales. The three unidimensional scales presented here have utility in their own right, however, for future investigations based on the factor analytic approach. Also, non-clinical scales developed in the future should be more than mere replications of the first two well defined factors.

Summary

Multiple correlations of three MMPI factor scales were found with the standard MMPI clinical scales and a large number of empirically derived MMPI scales in order to ascertain how well the MMPI domain could be represented by these three predictor scales. Factor scale M_1 seemed to be a member of the second-order dimension called Dynamic Integration *vs.* General Complaint and as such had the most prominent relationships with most of the other MMPI scales. Factor scale M_2 related with a second-order Repressive-Expressive dimension and was less well represented in the MMPI scales than was M_1. Factor scale M_3 had the least representation in the other MMPI scales.

The combined factor scales did a good job of predicting most of the non-clinical scales because the high saturation of the scale items with social desirability brings these scales into the sphere of the second-order Dynamic Integration *vs.* General Complaint dimension. For the standard nine clinical scales Hy, Pd, Mf, and Pa had an extremely low proportion of variance accounted for by the three predictors. In order to do an adequate job of replacing the clinical scales at least one or two additional factor scales would be necessary; the results indicated that most of the clinical scales not well predicted have a multidimensional complexity not accounted for completely by the three factor scales.

References

1. Barron, F. An ego strength scale which predicts response in psychotherapy. *J. consult. Psychol.*, 1953, *17*, 327-335.
2. Block, J. MMPI ego-control scale number 3. Personal communication, IPAR, Univ. Calif., Berkeley, 1959.
3. Cattell, R. B. Second-order personality factors in the questionnaire realm. *J. consult. Psychol.*, 1956, *20*, 411-418.
4. Cuadra, C. A. A psychometric investigation of central factors in psychological adjustment. Ph.D. dissertation, Univ. Calif., 1953.
5. Edwards, A. L. *The social desirability variable in personality assessment and research.* New York: Dryden, 1957.
6. Gocka, E. F. and Marks, J. B. Second-order factors and response sets in personality measurement. Unpublished Report, Research Service, Veterans Administration Hospital, American Lake, Washington, 1959.
7. Gough, H. G. A new dimension of status: I. Development of a personality scale. *Amer. social rev.*, 1948, *13*, 401-409.
8. Gough, H. G., McClosky, H., Meehl, P. E. A personality scale for measuring dominance. *J. abn. soc. Psychol.*, 1951, *46*, 360-366.
9. Gough, H. G., McClosky, H., Meehl, P. E. A personality scale for social responsibility. *J. abn. soc. Psychol.*, 1952, 47, 73-80.
10. Gough, H. G. Studies of social intolerance: II. A personality scale for anti-semitism. *J. soc. Psychol.*, 1951, *30*, 247-255.
11. Hanvik, L. J. MMPI profiles in patients with low-back pain. *J. consult. Psychol.*, 1951, *15*, 350-353.
12. Hathaway, S. R., and Briggs, P. F. Some normative data on new MMPI scales. *J. clin. Psychol.*, 1957, *13*, 364-368.
13. McClelland, W. A. A preliminary test of Role-Playing ability. *J. consult. Psychol.*, 1951, *15*, 102-108.
14. Mees, H. L. Preliminary steps in the construction of factor scales for the MMPI. Naval Research Contract Nonr-477(08) and Public Health Research Grant M-743(C3), Division of Counseling and Testing, Univ. of Wash., Seattle, July, 1959.
15. Navran, L. A rationally derived MMPI scale to measure dependency. *J. consult. Psychol.*, 1954, *18*, 192.
16. Rozynko, V. MMPI Internalization-Externalization scale number 3. Personal communication, Research Service, Veterans Administration Hospital, American Lake, Washington, 1959.
17. Siegel, S. M. The relationship of hostility to authoritarianism. *J. abn. soc. Psychol.*, 1956, *52*, 368-372.
18. Taylor, Janet A. A personality scale of manifest anxiety. *J. abn. soc. Psychol.*, 1953, *48*, 285-290.
19. Welsh, G. S. Factor dimensions A and R. In Welsh, G. S. and Dahlstrom, W. G. *Basic Readings on the MMPI in Psychology and Medicine:* Univ. of Minn. Press, 1956, 264-281.
20. Williams, H. L. The development of a caudality scale for the MMPI. *J. clin. Psychol.*, 1952, *8*, 293-297.

TWO MMPI SCALES FOR ALCOHOLISM: WHAT DO THEY MEASURE?[1]

MAURICE KORMAN

University of Texas Southwestern Medical School

PROBLEM

During an investigation on the definition of the alcoholism syndrome, it was noted that two empirical MMPI-based alcoholism scales were not substantially related ($r = .26; p = .05$). The AL scale, developed by Holmes[2] and a second scale devised by Hoyt and Sedlacek[4] (AHS) contain 59 and 68 items respectively, with an overlap of 28 items. Both scales were constructed through item analyses of the MMPI records of alcoholics and a portion of the original Minnesota standardization group, and both scales are reported to differentiate significantly alcoholics from normals.

Since the scales were constructed against the criterion of a diagnosis of alcoholism, they may be said to have parallel concurrent validity; their low correlation, however, implies that they may reflect different subsets of the universe of attributes which differentiate alcoholics from normal, i.e., have different construct validity[3]. This study attempts to investigate what the two scales "measure" by examining their relationship to a number of variables reflecting the physical, psychological and social adjustment of the alcoholic.

METHOD

The subjects were 61 patients who had been referred to an out-patient alcoholism clinic and who had undergone medical, psychiatric and psychological evaluation. Measures of the following dimensions or variables were considered in relation to AL and AHS:

1. *Scale of preoccupation with alcohol.* This ten item scale was assembled by Jackson[6] through modified Gutman scaling. It represents a sample of alcohol related behavior (e.g., sneaking drinks at parties) and is suggested to measure extent of alcoholism.

2. *Rapidity of development of "preoccupation" behavior.* The age of occurrence of the earliest reported behavior on the scale of preoccupation with alcohol was subtracted from the age at which the most recent behavior was initiated. This variable indicates the rapidity with which critical alcohol related behavior developed.

3. *Total medical rating.* Following laboratory tests and a physical examination, an internist[3] rated the involvement of each of the eight traditional physical systems on a 1 - 5 scale. These ratings were then summed for each patient.

4. *Alcohol medical rating.* Ratings reflecting the degree to which the patient's physical condition appears to be related to his consumption of alcohol were likewise performed by the same physician.

5. *Sum of MMPI clinical scales.* The *T*-scores of the eight basic clinical scales plus *Sie* were summed. The scales were used as continous variables in order to obtain a single overall score reflecting general psychological-psychiatric maladjustment.

6. *Number of MMPI scales above 70.* This merely consists of the number of clinical scales which reached or exceeded a *T*-score of 70. Like the previous variable, it was included as a general indicator of mental status.

7. *Strength of need for alcohol.* This locally devised scale is composed of 16 items of the type "I drink because . . . it makes me feel less shy." Under the hypothesis that such a scale represents a measure of the number of drive-reinforcement contingencies maintaining alcohol behavior, it is included as a measure of the pressure or drive for alcohol.

8. *Total years.* This represents the patient's estimate of the number of years that he has been engaged in both heavy and light drinking.

9. *Frequency index.* This variable, not independent from the previous one, represents a combined estimate of duration and frequency of drinking.

10. *Number of spells.* The patient's estimate of the number of drinking spells or binges which he has experienced in a year's time.

---

[1]This research was supported by a grant from the Hoblitzelle Foundation.
[3]Dr. Norman Carter's help is gratefully acknowledged.

11. *Social adjustment.* Patients were rated on an adaptation of the scale by Barabee et. al.,[1] which stressed the patient's adjustment in the areas of family, occupation and community roles.

12. *Intellectual functioning.* The sum of the vocabulary and the abstraction scores from the Shipley-Hartford Scale was used to reflect this dimension.

13. *Intellectual deterioration.* This was measured by the conceptual quotient as derived from the Shipley-Hartford Scale in the standard fashion.

AL and AHS scores had been obtained from the short form of the MMPI; as a result only 43 of the AL items and 52 of the AHS items were recorded and used in this study. Since it was not possible to obtain a full set of data on each patient, the *N* varies somewhat, the median *N* being 51.

RESULTS AND DISCUSSION

Table 1 reports the relationship between AL and AHS and the 13 variables listed above, together with an estimate of the significance of the difference between the two scales' power to account for the variance in each variable. Both scales appear to be materially related to other important dimensions in the assessment of alcoholics, as witnessed by the fact that a third of the correlations in table 1 are significant at the .05 level.

TABLE 1. RELATIONSHIP OF AL AND AHS TO OTHER VARIABLES (PRODUCT - MOMENT CORRELATIONS)

Variable	AL	AHS	p^1
1. Preoccupation scale	.17	.03	
2. Development of preoccupation	−.41**	−.07	
3 Total medical rating	.13	.26	
4. Alcohol medical rating	.08	.18	
5. Sum of MMPI scales	.34**	−.39**	<.001
6. No. of MMPI scales < 70	.31*	−.31*	<.001
7. Need for alcohol	.40**	−.15	<.001
8. Total years	.14	.20	
9. Frequency index	−.27	.33*	<.002
10. No. of spells	.13	.03	
11. Social adjustment	.12	−.19	
12. Intellectual functioning	.51**	.41**	
13. Intellectual Deterioration	−.05	.18	

[1]Significance of difference between *r*'s calculated through NcNemar's test[6, p. 148].

*$p < .05$
**$p < .01$

The two scales do indeed appear to be measuring different aspects of alcoholism. A high AL score is associated with a stronger drive for alcohol, greater severity of psychological symptomatology, and a relatively rapid development of critical alcohol related behavior. AHS, on the other hand, seems to reflect a greater frequency of alcohol consumption with, however, relatively less of a subjectively experienced pressure to drink, and a less acute psychological involvement. It is interesting to note that while both scales are essentially unrelated to total years of drinking, a high AHS suggests frequent bouts with alcohol during those years while a high AL points in the opposite direction. Both scales are found to be positively and significantly related to intelligence, suggesting that low scores on both scales may sometimes reflect unfamiliarity with or inability to handle paper and pencil type tests.

An attempt to inductively summarize the properties of the two scales into underlying constructs might posit AL as a measure of the acute inner upheaval of the alcoholic while AHS would be seen as reflecting the more stabilized process of chronic alcoholism. Of the two, AL seems to be embedded in a somewhat richer nomological network since it appears to be more distinctly related to other significant variables. Further research will attempt to evaluate the prognostic import of the relationship of AL to AHS.

Summary

This study investigated what two MMPI scales for alcoholism "measure" by examining their relationship to a number of variables reflecting the physical, psychological and social adjustment of alcoholics. The first scale (Al) appears to reflect a stronger drive for alcohol, more severe psychological maladjustment, and a relatively rapid development of critical alcohol related behavior. The second scale (Ahs) seems to be associated with less acute psychological symptomatology, a greater frequency of drinking behavior without, however, the implication of a correspondingly intense "need" to drink.

References

1. Barabee, P., Barabee, Edna L. and Finesinger, J. E. A normative social adjustment scale. *Am. J. Psychiat.*, 1955, *112*, 252-259.
2. Button, A. D. A study of alcoholics with the Minnesota Multiphasic Personality Inventory. *Quart. J. Stud. Alc.*, 1956, *17*, 263-281.
3. Cronbach, L. J. and Meehl, P. E. Construct validity in psychological tests. *Psychol. Bull.*, 1955, *52*, 281-302.
4. Hoyt, D. P. and Sedlacek, G. M. Differentiating alcoholics from normals and abnormals with the MMPI. *J. clin. Psychol.*, 1958, *14*, 69-74.
5. Jackson, Joan K. The definition and measurement of alcoholism. *Quart. J. Stud. Alc.*, 1957, *18*, 240-262.
6. McNemar, Q. *Psychological Statistics*. Wiley: New York, 1955.

MMPI HIGH POINT CODE FREQUENCIES IN A STATE HOSPITAL POPULATION[1]

Reuben J. Silver[2] and Lloyd K. Sines[2]

Fergus Falls (Minn.) State Hospital

Problem

Interpretation of Minnesota Multiphasic Personality Inventory (MMPI) profiles has continued to emphasize configural aspects rather than individual scale scores. Further, it has become increasingly popular to code profiles, employing one of the existing coding techniques[2, 4], and to ascertain empirically the behavioral characteristics associated with specific code types in a particular clinical setting or population. Such "actuarial" treatment of MMPI data is facilitated by a determination of the distribution of code types in specific populations. That particular populations have different code frequencies has been reported by Hathaway and Meehl[3] for normals and for a medical school psychiatric inpatient population, and by Guthrie[1] for a medical outpatient sample. The present study reports the distribution of MMPI codes in three groups within a state mental hospital population, viz, recent admissions, releases, and continued treatment groups.

Method

MMPIs were obtained from the following three groups of patients at the Fergus Falls State Hospital:

Admission group. Beginning February, 1958, all newly admitted patients under age 65 were routinely administered the MMPI if they were literate and judged amenable to psychological testing. Admission MMPI profiles obtained from a randomly selected sample of 200 male and 200 female patients were selected for consideration.

---

[1]Supported in part by a grant from the Minnesota Department of Public Welfare.

[2]Now at the University of Minnesota.

Continued treatment group. Patients on the continued treatment wards, exclusive of the geriatric and medical units, who were able to read and to cooperate in testing, were given the MMPI. Profiles from a random sample of 200 males and 200 females were selected for analysis.

Releases. All patients who received an MMPI as a part of the routine admission procedure were retested prior to release from the hospital. The first 100 males and 100 females so tested constitute the release group.

Each MMPI was coded using Welch's method[4]. In grouping the MMPIs by high point codes, it was decided to consider the two digits as interchangeable, e.g., code 21 was grouped with code 12. It should be noted that the two digit codes are comprised of the eight clinical scales only, since scale 5 (Mf) was not included.

Results and Discussion

The incidence of the MMPI two digit codes is shown in Table 1. It should be noted that the data for the release group are reported as independent of their respective admission codes, *i.e.*, from the data presented it is not possible to determine admission codes for particular release cases. It can be seen (Table 2) that seven code types account for almost one half of the cases in all three groups. However, certain code types are more frequent in some groups than in others. A 68 code, for example, is most characteristic of the continued treatment patients. On the other hand, a 34 code is most typical of male patients at time of release. (Apparently females who are released are more heterogeneous, with respect to two digit code, than are males). Finally, codes 23 and 27 seem to be most typical of recently admitted patients. Although generalizations from these data must be made with caution, it appears that 68 codes may be associated with chronicity and that neurotic features predominate in newly admitted cases.

Table 1. Frequency of MMPI High Point Codes Among Groups of State Mental Hospital Patients

Two digit code	Admissions		Release		Continued Treatment	
	Male (N = 200)	Female (N = 200)	Male (N = 100)	Female (N = 100)	Male (N = 200)	Female (N = 200)
12/21	16	10	4	4	7	2
13/31	7	4	3	3	8	8
14/41	5	1	3	5	8	3
16/61	0	1	2	1	1	1
17/71	1	0	3	0	1	2
18/81	7	4	1	1	16	3
19/91	1	1	0	1	0	0
23/32	18	23	4	8	6	5
24/42	17	12	9	5	8	8
26/62	7	10	1	5	7	6
27/72	20	15	9	4	4	5
28/82	6	3	3	2	13	8
29/92	2	1	0	1	2	1
34/43	11	10	20	9	4	11
36/63	2	4	1	5	2	3
37/73	1	0	1	2	0	0
38/83	1	5	1	1	3	2
39/93	0	1	1	2	1	1
46/64	10	18	5	6	9	23
47/74	6	3	4	5	0	1
48/84	14	16	14	6	14	20
49/94	12	11	7	9	6	5
67/76	4	4	1	2	1	0
68/86	10	23	1	4	43	55
69/96	2	7	1	4	9	11
78/87	13	6	0	0	9	4
79/97	0	2	0	0	0	0
89/98	7	5	1	5	18	12

TABLE 2. PERCENTAGE OF PATIENTS IN THE THREE STATE HOSPITAL SAMPLES IN EACH OF THE SEVEN MOST COMMON TWO DIGIT CODES

Two digit code	Admission		Release		Continued Treatment	
	Male (N = 200)	Female (N = 200)	Male (N = 100)	Female (N = 100)	Male (N = 200)	Female (N = 200)
23	9.0	11.5	4.0	8.0	3.0	2.5
24	8.5	6.0	9.0	5.0	4.0	4.0
27	10.0	7.5	9.0	4.0	2.0	2.5
34	5.5	5.0	20.0	9.0	2.0	5.5
46	5.0	9.0	5.0	6.0	4.5	11.5
48	7.0	8.0	14.0	6.0	7.0	10.0
68	5.0	11.5	1.0	4.0	21.5	27.5
Total	50.0	58.5	62.0	42.0	44.0	63.5

Although all possible two digit codes are found in the population studied, some rarely occur. Codes 16, 17, 19, 29, 37, 38, and 67, which occur infrequently in this study, are also rare in other populations[1, 3]. Because of the infrequency of these codes, it will be extremely difficult to determine behavioral correlates of such codes via actuarial techniques, since such methods require sizable samples for the determination of stable population characteristics.

SUMMARY

The present study reports the distribution of MMPI code types in three groups within a state hospital population: 200 male and 200 female recent admissions, 200 male and 200 female continued treatment patients, and 100 male and 100 female patients upon discharge. Seven code types accounted for almost one half of the patients in all groups. It was also suggested that certain two digit codes occur rarely regardless of the population studied, thus limiting possibilities for the actuarial determination of stable base-rate population characteristics.

REFERENCES

1. GUTHRIE, G. M. Common characteristics associated with frequent MMPI profile types. In G. S. Welsh and W. G. Dahlstrom (Ed.), *Basic readings on the MMPI in psychology and medicine.* Minneapolis: Univer. of Minn. Press, 1956, 145-150.
2. HATHAWAY, S. R. A coding system for MMPI profile classification. In G. S. Welsh and W. G. Dahlstrom (Ed.), *Basic readings on the MMPI in psychology and medicine.* Minneapolis: Univer. of Minn. Press, 1956, 126-129.
3. HATHAWAY, S. R., and MEEHL, P. E. Psychiatric implications of code types. In G. S. Welsh and W. S. Dahlstrom (Ed.), *Basic readings on the MMPI in psychology and medicine.* Minneapolis: Univ. of Minn. Press, 1956, 136-144.
4. WELSH, G. S. An extension of Hathaway's MMPI profile coding system. In G. S. Welsh and W. S. Dahlstrom (Ed.), *Basic readings on the MMPI in psychology and medicine.* Minneapolis: Univ. of Minn. Press, 1956, 130-132.

A SHORTENED SCHIZOPHRENIC SCALE: APPLICATION TO CONFINED GROUPS[1]

RALPH ROBINOWITZ

Veterans Administration Hospital
Waco, Texas

In a recent article, Clark and Danielson[1] presented a shortened scale for detecting schizophrenia. The scale has similarity to others widely used for induction screening by the military, and is based on 20 items from the Sc scale of the Minnesota Multiphasic Personality Inventory (MMPI) found "most valid in discriminating between 'normals' and diagnosed schizophrenics." Attention has been called by Cuadra and Reed[2], to the difficulty in detecting many cases of schizophrenia through short psychological measures, such as the Sc scale of the MMPI. One problem noted is the similarity of response of diagnosed schizophrenics and persons with similar dynamics, who might also feel "alienated, misunderstood, and out of place".

The present study compares the Clark-Danielson scale with similar groups of schizophrenics and non-schizophrenics. Confined military offenders show many of the same feelings as hospitalized schizophrenics. Would it not be expected that the responses to scale items of the military offenders would be similar to those of the schizophrenic group? The question also arises as to the meaning of confinement in both groups. Does confinement in and of itself produce schizophrenic-like responses in a non-psychotic confined group? What part of the measured responses of schizophrenics to the scale is a factor of confinement alone?

Method

Two hundred male subjects, all below thirty-five years of age, were administered the shortened schizophrenic scale. One hundred subjects were recent military veterans on their first admission to a large neuropsychiatric hospital. All were seen within a two week period following admission, were from a closed ward and had been diagnosed as schizophrenic by at least two psychiatrists. The second hundred subjects in the present study were servicemen convicted of military offenses and confined to a post stockade at a major Army installation. All were judged nonpsychotic by a psychiatric interview. The scale was administered within a few days following admission to the stockade.

Results

The responses of each member of the two groups were scored to permit computation of the number of "maladjusted responses"[1]. The mean number of maladjusted responses for the schizophrenic group was 7.11, S. D., 4.28. In the nonschizophrenic group (military offenders), the mean number of maladjusted responses was 3.86, S. D. 3.80. The mean number of maladjusted responses of the two groups was found to differ significantly ($t = 5.87$, significant at the .01 level). The schizophrenic group defined for the present study was found to be quite similar to the schizophrenic group studied by Clark and Danielson[1]. Not only was there close agreement in the two groups as to mean scores (number) of maladjusted responses, but the standard deviations of such scores were almost identical.

The range of number of maladjusted responses for the two groups, however, is almost the same (0 - 18 for the schizophrenic group, 0 - 16 for the military delinquent group). Despite the general effectiveness of the scale in group differentiation of schizophrenics, there is no effective means of its utilization other than as a screening instrument[1].

[1]Appreciation is expressed to associates of the author on the staffs of the Post Stockade and U. S. Army Hospital, Fort Gordon, Georgia, and the Veterans Administration Hospital, Waco, Texas; and to Dr. Wayne H. Holtzman, who serves as Research Consultant and Mrs. Doris K. Neumann, Research Assistant, at the latter hospital. Responsibility for statements and conclusions remains that of the author.

When differences in responses to the individual items comprising the "shortened schizophrenic scale" were studied for the two groups, significant differences (at the .01 level) were found for only eight of the twenty items. MMPI item numbers 194, 334, 212, 266, 333, 323, 182, and 301 were responded to differently by the two groups. Several items on the shortened scale as expected did not differentiate between the two groups. Though items involving recognition of mental disturbance and unusual sensory and emotional experiences do usually differentiate the groups, it is only the least subtle of such items that are discriminatory. Less direct items even where clearly apparent, such as those involving memory function, did not significantly differentiate the group. However, those items involving feelings of hostility towards people and of being victimized by society and alienated by people were answered similarly by the schizophrenics and the military delinquents. It is of interest that even where special conditions rendered twelve of the items noneffective, the shortened schizophrenic scale did significantly differentiate schizophrenics from a group of military delinquents.

Summary

The shortened schizophrenic scale devised by Clark and Danielson was studied with a hospitalized schizophrenic group and a confined military delinquent group to study the effect of confinement on the scale. The scale was found to differentiate between the two groups with regard to mean number of maladjusted responses. Even though several items were found nondiscriminatory, the overall effectiveness of the scale for rapid screening purposes was supported. However, the overlap in range of maladjusted response score for the two groups again points out the danger in using the scale for individual diagnosis. The military delinquents as a group responded to the shortened scale more in the manner of normals than of schizophrenics, showing delinquent behavior to be unlike schizophrenia with regard to responses to items on the screening scale.

References

1. Clark, J. H. and Danielson, J. R. A shortened schizophrenic scale for use in rapid screening. *J. soc. Psychol.*, 1956, *43*, 187-190.
2. Cuadra, C. A. and Reed, C. F. An introduction to the Minnesota Multiphasic Personality Inventory. Unpublished Handbook (mimeographed) based on lectures read at the Veterans Administration Hospital, Downey, Illinois, Spring 1954.
3. Hathaway, S. R. and McKinley, C. *The Minnesota Multiphasic Personality Inventory.* Minneapolis: Univer. Minnesota Press, 1943.

MMPI PROFILES OF PSYCHOPATHS AND PRISONERS

LEONARD KINGSLEY[1]

Mental Hygiene Consultation Service, Ft. Meade, Md.

This study was designed to test the following hypothesis: Are MMPI scales, with particular reference to *Pd* and Welsh's Anxiety Index (AI) and Internalization Ratio (IR), significantly higher in psychopaths than in a matched group of nonpsychopathic prisoners or non-incarcerated "normal" controls? The unique feature of the current study was the individual matching of the subjects.

Method

There were three groups of subjects in this study: 25 psychopathic personalities, 25 nonpsychopathic offenders and 50 non-incarcerated controls. All subjects were enlisted men in the military service. The 25 psychopaths were diagnosed as either Antisocial or Asocial Personalities (2, p. 9) by army psychiatrists. The 25 nonpsychopathic prisoners were considered to have "no psychiatric disease" (2, p. 18) by the same psychiatrists. The controls were 50 enlisted men, on duty at a military in-

stallation, whose records gave no history of crimes, save for minor traffic violations. The two prisoner groups were matched man-for-man for education (within two grades) and for age (within two years); 89% of the subjects met the criterion for age and 92% for education. Subsequently, the two prisoner groups were combined into a general prisoner category and the offenders matched individually with the controls for these two variables. Ninety-six per cent met the criterion for age and for education.

Results

The means and standard deviations of the MMPI scales of the three groups are given in Table 1. All scales were within the average (50-70) range, with *Pd* as the highest in all three groups. The *t* values for correlated means indicate that none of

Table 1. Means and Standard Deviations of MMPI Scores of Psychopaths, Nonpsychopathic Prisoners, and Normal Controls

Scales	Psychopaths		Nonpsychopathic prisoners		Normal controls	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
L	51.28	8.19	52.36	9.08	52.10	4.04
F	61.52	15.05	62.92	12.25	52.86	6.85
K	53.04	9.09	53.24	10.60	54.66	9.85
Hs	57.60	15.15	60.48	16.13	52.08	8.05
D	58.60	14.56	60.60	17.69	54.60	10.51
Hy	60.44	9.66	60.52	10.45	54.54	7.56
Pd	67.52	6.65	69.28	8.75	60.08	10.37
Mf	55.28	8.29	53.04	6.92	52.78	7.17
Pa	59.28	10.48	56.96	12.87	51.06	9.53
Pt	59.92	14.78	60.28	12.87	53.34	9.76
Sc	63.84	15.23	63.92	16.05	54.58	6.93
Ma	62.96	10.06	62.52	10.61	59.96	10.30
AI	58.96	22.92	59.96	24.26	59.53	19.00
IR	.92	.14	.97	.24	.93	.13

the MMPI scales were significantly greater in one offender group as compared to the other. Because of the similarity of the results, both groups were then combined into a general prisoner category and compared with the controls. Seven of the clinical scales (*F*, *Hs*, *Hy*, *Pd*, *Pa*, *Pt*, and *Sc*) were significant at the .01 level of confidence with *Pd* having the highest *t* value.

To measure the elevation of *Pd* from the rest of the profile,[1] the sum of the other clinical scales, excluding *Mf*, were averaged and compared individually with the *Pd* score of each subject. *Mf* was excluded because it was considered to be more of an interest scale than a clinical one. The *t*s for the difference between the correlated means of *Pd* and the sum of the other scales were 4.60, 5.28, and 4.38 for psychopaths, prisoners, and controls, respectively—beyond the .01 level in all three groups. *Pd* was thus significantly elevated from the rest of the profile in all three. In view of the high degree of similarity between these scores, no *t*s for psychopaths vs. nonpsychopaths or general prisoners vs. controls were done; it is fairly evident that no significant differences exist between the groups in this respect.

The AI and IR were also computed individually for all subjects; means and standard deviations are given in Table 1. The *t* values for the correlated means of psychopaths vs. nonpsychopaths and general prisoners vs. controls do not indicate that these ratios were significantly greater in one group of either pair as compared to the other.

Summary

Twenty-five psychopaths were matched individually for age and education with 25 prisoners with "no psychiatric disease." There were no significant differences between the two groups on any of the MMPI scales, the Anxiety Index, the

Internalization Ratio, or the elevation of *Pd* from the mean of seven of the other MMPI scales. Because of the similarity of the results, both prisoner groups were combined into a general category and matched individually for age and education with 50 enlisted men who gave no history of crimes. Seven of the scales significantly differentiated these two groups. The AI and IR did not, nor did the elevation of *Pd* in relation to the other clinical scales. The prisoner population studied may be described as possessing hostile tendencies, somatic complaints, paranoid tendencies, ruminative worries and as socially withdrawn.

References

1. McKinley, J. C. and Hathaway, S. R. The Minnesota Multiphasic Personality Inventory. V. Hysteria, Hypomania and Psychopathic Deviate. *J. appl. Psychol.*, 1944, *28*, 153-174.
2. *Joint Armed Forces Nomenclature and Methods of Recording Psychiatric Conditions.* Department of the Army, Bureau of Medicine and Surgery, U. S. Navy, Department of the Air Force, Washington 25, D. C. June 1949. Pp. 20.

COMPARISON OF MMPI PROFILE DIFFERENCES AMONG NEGRO-WHITE ADOLESCENTS

JOHN C. BALL

University of Kentucky

Problem and Method

This study compares Minnesota Multiphasic Personality Inventory data from Negro and white ninth grade students in two Kentucky integrated public schools. The subjects were 200 ninth grade students of whom 31 were Negro and 169 white. These adolescents constituted the entire ninth grade in attendance at the time the MMPI was administered in two Central Kentucky towns, except that 24 subjects who completed invalid profiles were eliminated from the comparison. The Negro and white students attended the same classes and resided in the same school districts.

In Table 1 the age, IQ, home composition, socio-economic status, mean grade achieved for previous semester, and educational retardation are presented. The Negroes and whites did not differ significantly in age. The mean IQ of the Negro students is some ten points lower than that of the white students. These tests were not, however, administered as a part of this study. The IQ comparisons are, therefore, based upon incomplete and incomparable scores and should be regarded only as indicative findings. For this reason no test of significance was employed. A greater percentage of Negroes came from broken homes than did the white; the difference was statistically significant between the two groups of boys, but not between the girls. With respect to socio-economic status, the Negroes were predominantly from the low-

Table 1. Age, IQ, Family Composition, School Grades, Educational Retardation and Socio-Economic Status of Negro and White Students by Sex

Subjects	N	Mean Age	Mean IQ	Broken Homes, Percent	Lower-Class, Percent	Mean Grade	Retarded Percent
Negro Boys	14	15.3	96.4	35.7*	85.8**	2.9**(D+)	50.0
White Boys	81	15.0	104.5	8.6	18.5	4.7 (C+)	31.2
Negro Girls	17	14.4	93.6	29.4	64.7**	4.1** (C)	11.8
White Girls	88	14.8	105.3	18.2	14.5	5.3 (B−)	18.2

Mean difference between Negro and white boys (or girls) significant at: *P < .05, **P < .01. No test of significance applied to IQ data, for reason see text.

est two of the seven classes on the Minnesota Parental Occupation Scale, while the whites mostly were from higher status families. Educational differences between the Negro and white ninth graders were evident with respect to present scholastic achievement, but not with respect to educational retardation. Thus, the Negro students had grade averages below those of the white students, but in number retarded one or more years in grade placement the two groups did not differ significantly.

Results

Valid MMPI profiles were completed by 200 of the 224 ninth graders tested. Criteria of validity were: ? < 100, L < 10, F < 16, and F-K < 9. Of the 24 students with invalid profiles, 11 were Negro and 13 white. Thus, 7.1 percent of the white and 26.2 percent of the Negro students had invalid profiles.

Of those with valid profiles, it may be noted from Table 2 that the Negro boys differ from the whites only on the Hs scale which is higher among the Negroes. The

Table 2. MMPI Mean Scale Scores and *t* Values of Negro and White Public School Students

Scale	Boys			Girls		
	Negro N = 14	White N = 81	*t*	Negro N = 17	White N = 88	*t*
L	4.4	3.4	1.67	3.3	4.2	1.61
F	7.4	5.7	1.66	8.8	5.2	4.13**
K	12.6	11.8	0.63	9.1	13.7	3.85**
1 - HS	56.4	50.4	2.20*	49.3	49.5	0.08
2 - D	58.1	53.3	1.44	52.2	50.6	0.73
3 - Hy	52.7	52.8	0.05	47.3	53.8	2.82**
4 - Pd	60.9	59.5	0.54	61.2	61.0	0.05
5 - Mf	53.7	53.5	0.08	53.5	51.2	0.99
6 - Pa	54.1	54.3	0.05	57.8	55.7	0.82
7 - Pt	60.1	58.0	0.70	57.5	57.2	0.12
8 - Sc	61.2	59.0	0.76	66.2	58.6	3.13**
9 - Ma	63.2	62.1	0.35	60.1	60.4	0.12
0 - Si	53.4	53.6	0.08	59.5	51.7	3.82**

*Validity scales are given in mean raw scores, clinical scales in mean T-scores. K corrections have been applied to scales 1, 4, 7, 8 and 9. Significance of difference between means: *P < .05, **P < .01.

Negro girls were higher on scales Sc, Si and F and lower on Pd and Ma than the white girls.

The percentage of individuals in the four groups—Negro boys, white boys, Negro girls and white girls—with one or more scales at or above T-score 70 was, respectively, 57.2, 54.3, 47.1 and 42.0; the differences were not significant.

Table 3. First Scale in MMPI Profile Among Negro and White Students by Sex

Highest MMPI Scales	Percent of Negro Students		Percent of White Students		Total No.
	Boys (N = 14)	Girls (N = 17)	Boys (N = 81)	Girls (N = 88)	
1. Hs, D, Hy, Pt	50.0	5.9	23.5	18.2	43
2. Pd, Ma	42.9	29.4	48.1	53.4	97
3. Pa, Sc	7.1	41.2	13.6	15.9	33
4. Mf, Si		23.5	14.8	12.5	27

Difference in MMPI Scale Types between Negro and White Students:

Boys: $\chi^2 = 5.616$, $P < .20$

Girls: $\chi^2 = 8.744$, $P < .05$

In Table 3 the Negro and white groups are compared with regard to the first or highest MMPI scale in the individual profiles. Haertzen and Hill's[2] classification of scale types in a somewhat revised form is used. Group 1 (Hs, D, Hy and Pt) is the neurotic classification. In group 2 (Pd and Ma), psychopathic tendencies (or hyperactivity in an adolescent sample) are indicated. In group 3 (Pa and Sc), schizoid tendencies are prominent (or in this normal group of adolescents, withdrawal and sensitivity). Group 4 (Mf and Si) is unclassified, although it is in some respects quite similar to group 3.

The Negro and white groups differ considerably with respect to personality configuration, as measured by the first scale in MMPI profiles. Neurotic tendencies (Hs and D) are most marked among the Negro boys, while withdrawal and introversion are characteristic of the Negro girls. The Si scale was especially notable in the profiles of the Negro girls (highest in 23.5 percent of the cases), while among the Negro boys the D scale was prominent (first in 28.6 percent of the profiles).

Discussion

More pronounced personality differences were found between the Negro and white girls than between the boys. The Negro girls were more withdrawn and introverted than the white girls; they also appeared to be unusually candid in their test-taking attitudes as seen from low L and K scores and high F scores. The higher Hy scores among the white girls appears to reflect socio-economic differences between the two groups.

The Negro and white students differ markedly on socio-economic status and to a lesser extent on the prevalence of broken homes. The question arises as to the effect of these factors upon the personality differences reported. Previous research[4] has indicated that lower class students score higher on the Pd and D scales and lower on Hy, Mf and Ma scales. Since the Negro and white subjects in the present study did not differ on Pd, Mf or Ma it is maintained that socio-economic factors do not account for the principal personality differences found. An exception to this might well be the Hy scale, which was higher among the white girls. The higher scores of the Negro boys on the D scale were associated with the high incidence of broken homes in this group rather than with status differences between Negro and white students.

The more frequent completion of invalid profiles by the Negro students was associated with poor academic performance and low IQ. A separate comparison of the 24 students who completed invalid profiles with the valid group revealed marked differences between the two on academic achievement and IQ; academic deficiency and low IQ were associated with invalid MMPI profiles. A similar finding concerning the association of high Cannot Say scores with below average grades and low IQ has been reported.[1]

It is pertinent to note that scale order or configuration of the MMPI was more discriminative than profile elevation itself. Thus, analysis of the highest scale in the profiles provided a more accurate delineation of personality variables than mere consideration of elevation. This is not surprising when it is recognized that profile elevation contrasts the individual's scores against established norms without regard to the relative importance of such elevation to the person himself. Conversely, scale position or code analysis provides a means of analyzing the *relative* importance of specific scales to the individual without regard to predetermined norms. The present findings then, support the emphasis given to the presentation and analysis of code types and profile configurations.[3]

Summary

A comparison of Negro and white ninth grade high school students with the MMPI revealed a high incidence of neurotic tendencies among the Negro boys as well as withdrawal and introversion among the Negro girls. The effect of age, IQ, broken homes, socio-economic status, academic achievement and educational retardation upon personality differences was considered.

References

1. Ball, John C. and Carroll, Donna. Analysis of MMPI Cannot Say Scores in an Adolescent Population. *J. clin. Psychol.*, 1960, *16*, 30-31.
2. Haertzen, Charles A. and Hill, Harris E. Effects of Morphine and Pentobarbital on Differential MMPI Profiles. *J. clin. Psychol.*, 1959, *15*, 434-437.
3. Hathaway, S. R. and Meehl, P. E. "Psychiatric Implications of Code Types," reprinted in *Basic Readings on the MMPI in Psychology and Medicine*. G. S. Welsh and W. G. Dahlstrom. (Eds.) Minneapolis: University of Minnesota Press, 1956.
4. Hathaway, S. R. and Monachesi, E. D. *Analyzing and Predicting Juvenile Delinquency with the MMPI*. Minneapolis: University of Minnesota Press, 1953.
5. Hokanson, Jack E. and Calden, George. Negro-White Differences on the MMPI. *J. clin. Psychol.*, 1960, *16*, 32-33.
6. Stanton, John M. Group Personality Profile Related to Aspects of Antisocial Behavior. *J. Crim. Law, Crimin. Pol. Sc.*, 1956, *47*, 340-346.

MMPI PERFORMANCE OF APHASIC AND NONAPHASIC BRAIN-DAMAGED PATIENTS*

DONALD G. DOEHRING AND RALPH M. REITAN

Indiana University Medical Center

Problem

Brain-damaged individuals with aphasic symptoms tend to exhibit more severe impairment of verbal abilities than do nonaphasic brain-damaged individuals[2, 3]. According to Wepman[6], aphasic symptoms are also accompanied by a large number of nonlanguage deviations. Among the 34 nonlanguage deviations that he lists are feelings of inadequacy, egocentricity, increased irritability and fatigability, euphoria, social withdrawal and seclusiveness, regressive behavior, anxiety and tension, and changing personality profile. Such affective disturbances could be a secondary result of the language deficit associated with aphasia; or these behavior deviations might be characteristic of brain-damaged individuals in general, regardless of the presence or absence of aphasic symptoms. Previous investigations by the second writer[4, 5] have shown that there were no substantial differences in the Rorschach Test performance of aphasic and nonaphasic patients, and that a group of brain-damaged patients undifferentiated with respect to aphasic symptoms exhibited pronounced neurotic symptoms in their performance on the MMPI. The present investigation involved a comparison of the MMPI performance of a group of brain-damaged patients with aphasic symptoms, a group of brain-damaged patients with no aphasic symptoms, and a group of hospitalized control patients. The purpose of this study was to obtain further information concerning the question of whether affective disturbances exhibited by aphasic individuals specifically result from the language deficit associated with aphasic symptoms, or whether these disturbances are also characteristic of brain-damaged individuals with no aphasic symptoms.

Method

The individual form of the MMPI was used in this study. Subjects were 51 white males, including 17 brain-damaged patients with aphasic symptoms, 17 brain-damaged patients with no aphasic symptoms, and 17 non-brain-damaged patients

*This study was supported by Grant B-1468 from the National Institutes of Health, Division of Neurological Diseases and Blindness. The writers wish to thank Mrs. Sharon Schumacher for her assistance in this study.

with neurotic symptoms. Presence or absence of aphasic symptoms was determined by an evaluation of patients' performance on a modified form of the Halstead-Wepman Screening Test. The three groups were carefully matched in triads with respect to age and number of years of education. The mean and SD of age and education are shown in Table 1. The two brain-damaged groups were also matched

TABLE 1. COMPARISON OF MMPI SCORES, AGE, AND EDUCATION OF BRAIN-DAMAGED PATIENTS WITH APHASIC SYMPTOMS (N = 17), BRAIN-DAMAGED PATIENTS WITH NO APHASIC SYMPTOMS (N = 17), AND A NON-BRAIN-DAMAGED CONTROL GROUP WITH NEUROTIC SYMPTOMATOLOGY (N = 17).

Variable	Aphasic Mean	Aphasic SD	Nonaphasic Mean	Nonaphasic SD	Control Mean	Control SD	F
Age	34.00	9.43	34.65	10.44	33.71	10.09	
Educ. (Yrs.)	10.00	2.91	10.59	2.65	10.82	3.16	
Hs + .5K	64.17	19.62	63.58	15.32	73.94	16.76	1.93
D	69.17	16.49	67.17	13.08	73.47	14.73	< 1
Hy	62.29	18.17	61.82	11.04	65.70	11.27	< 1
Pd + .4K	65.23	6.66	57.76	11.67	64.00	14.41	2.14
Mf	54.76	9.72	53.41	10.07	59.41	9.37	1.81
Pa	61.94	14.99	55.58	10.65	57.17	9.20	2.77
Pt + 1K	66.53	16.64	61.11	14.35	68.29	14.66	1.03
Sc + 1K	74.00	17.32	60.41	18.00	69.64	18.68	2.44
Ma + .2K	60.17	15.08	55.11	9.16	56.29	13.24	< 1
L	5.47	2.44	4.76	1.50	3.23	1.41	5.77*
F	9.47	9.69	4.17	3.39	5.53	3.46	3.18
K	54.68	11.09	54.57	9.22	50.04	9.70	< 1
?	11.29	10.39	12.11	9.90	10.88	9.95	< 1
A	18.87	11.26	17.23	7.48	20.82	10.58	< 1

*Significant at the .01 level.

in terms of type of brain lesion. The aphasic group contained three patients with intrinsic cerebral tumors, seven with traumatic head injuries, two with cerebral vascular disease, one with multiple sclerosis, one with cerebral abscess, one with subdural hematoma, one with chronic brain syndrome, and one with excision of the left temporal lobe for relief of epilepsy. The nonaphasic group contained three patients with intrinsic cerebral tumors, eight with traumatic head injuries, two with cerebral vascular disease, two with multiple sclerosis, one with chronic brain syndrome, and one with excision of the left temporal lobe for relief of epilepsy. Since a previous study by Reitan[5] indicated that brain-damaged patients tend to have elevated scores on the "neurotic triad" of the MMPI, the control group for this study was selected to include subjects with definite neurotic symptomatology. All subjects in the control group had negative neurological and anamnestic findings with regard to brain damage.

RESULTS AND DISCUSSION

Analysis of variance was used for the evaluation of differences among mean scores on the MMPI scales, including the Taylor Anxiety Scale (A). Table 1 shows means and SDs of the scores for each group, and *F* ratios derived from the between-groups comparison of the analysis of variance for each scale. Raw scores were used for L, F, ?, and A, and T scores were used for the remaining scales. Of the 14 comparisons, only the difference among group means on the L scale was statistically significant. Further analysis of the differences among L scale means by the New Multiple Range Test[1] revealed that the control group made significantly fewer L responses then either the aphasic group (.01 level) or the nonaphasic group (.05 level). The aphasic and non-aphasic groups did not differ significantly in number of L responses.

The brain-damaged patients with aphasic symptoms did not differ appreciably from the brain-damaged patients with no aphasic symptoms in their performance on the MMPI clinical scales and on the Taylor Anxiety Scale. Nor did the two

brain-damaged groups differ markedly from the group of controls with neurotic symptomatology in their performance on these scales. The results of this study support Wepman's contention that aphasic patients tend to manifest personality disturbances in addition to the language disturbance directly associated with aphasic symptoms. However, the similarity of the performance of the nonaphasic brain-damaged group to that of the aphasic brain-damaged group suggests that the personality disturbances described by Wepman are characteristic of brain-damaged patients in general, irrespective of the presence or absence of aphasic symptoms.

While the similarity of the three groups would seem to be the more important point, the relatively large number of L responses made by the brain-damaged groups can be interpreted as an indication that the clinical scale scores for these groups may be attenuated by attempts to choose socially acceptable responses. If anything, then, the high L scale scores of the brain-damaged groups suggest that the clinical scale scores underestimate the extent of affective disturbance in these groups. It is beyond the scope of this study to explain why brain-damaged patients, who are usually considered to be rather apathetic, should be more highly motivated to make socially acceptable responses than are non-brain-damaged patients with pronounced neurotic symptomatology.

Summary

The individual form of the MMPI was given to 51 white males, including 17 aphasic brain-damaged patients, 17 nonaphasic brain-damaged patients, and 17 non-brain-damaged patients with neurotic symptomatology. All three groups were matched in age and education, and the two brain-damaged groups were matched in terms of type of brain lesion. There were no significant differences among group means for the clinical scales of the MMPI or the Taylor Anxiety Scale. However, both brain-damaged groups made significantly more L scale responses than the control group. It was concluded that both aphasic and nonaphasic brain-damaged patients manifest personality disturbances similar to those of non-brain-damaged patients with neurotic symptomatology.

References

1. Duncan, D. B. Multiple range and multiple F tests. *Biometrics*, 1955, *11*, 1-42.
2. Klove, H. and Reitan, R. M. Effects of dysphasia and spatial distortion on Wechsler-Bellevue results. *AMA Arch. Neurol. Psychiat.*, 1958, *80*, 708-713.
3. Reitan, R. M. Intelligence and language functions in dysphasic patients. *Dis. nerv. Systm.*, 1954, *15*, 131-137.
4. Reitan, R. M. The performance of aphasic, non-aphasic and control subjects on the Rorschach Test. *J. gen. Psychol.*, 1954, *51*, 199-212.
5. Reitan, R. M. Affective disturbances in brain-damaged patients. *AMA Arch. Neurol. Psychiat.*, 1955, *73*, 530-532.
6. Wepman, J. M. *Recovery from aphasia.* New York: Ronald Press, 1951.

TWO TYPES OF PARANOID SCHIZOPHRENIA[1]

BENJAMIN KLEINMUNTZ

Carnegie Institute of Technology

PROBLEM

In studies by Guertin and Zilaitis[1] and Guertin[2], four types of paranoid schizophrenics were disclosed by factor analysis and subsequently cross-validated. The present study substitutes performance on the semantic differential technique for the MMPI and Bender-Gestalt, and limits itself to the investigation of only two types of paranoid schizophrenia.

PROCEDURE

A group of 50 clinically homogeneous paranoid schizophrenics was selected from the neuro-psychiatric ward of the Veterans Administration Hospital at Minneapolis. The psychotic group was rated by five staff psychiatrists and psychologists and was considered homogeneous with respect to delusions, hallucinations and severity of symptoms. On the basis of MMPI's routinely administered on admission, the psychotic group was then dichotomized into 25 MMPI *Pa* scale test misses and 25 MMPI *Pa* scale test hits. There were also 25 normals selected from the general medicine and surgery ward, and all three groups were comparable on the variables of age, intelligence, education and motivational level.

The Semantic Differential[3, 4, 5] was used to compare the psychotic subgroups with themselves and with the normals. The Ss were asked to rate a total of ten concepts on eleven scales. Three concepts were included to represent the self (ME) and others (THEY and PEOPLE); and three concepts were arbitrarily chosen because they represent environmental objects which on the basis of clinical judgment seem especially meaningful to paranoid psychotics (FRIENDS, POISON and THE MIND). Of the remaining four concepts, two were MMPI, *Pa* Obvious items[6] which were converted into rateable form (A PERSON WHO IS SURE HE GETS A RAW DEAL FROM LIFE and A PERSON WHO BELIEVES HE IS CONDEMNED); and two were *Pa* Subtle items (PEOPLE WHO WOULD LIKE TO GET AHEAD and PEOPLE WHO TELL LIES TO KEEP OUT OF TROUBLE).

RESULTS

The semantic structures of the three groups were plotted in three dimensional space as illustrated in Figures 1, 2, and 3. The spatial orientation for the three semantic structures is that "mentally healthy" is down, and "mentally ill" is up; "weak" is to the left, and "strong" is to the right; "excitable" is away from the viewer, and "calm" is toward the viewer. Using this orientation, the reader may observe the following aspects in the three structures:

1. A comparison of the three figures indicates that generally the cluster of concepts for test misses is more similar to that of normals than to that of test hits. For example, concepts C, E, H and I are inter-related and are a separate cluster for test misses and normals, whereas for test hits these concepts are related also to concepts A, G and J.

2. In Figure 2 it is illustrated that test hits described themselves and the Mind as mentally ill, weak and excitable. Test misses, on the other hand, described themselves and the Mind as mentally healthy, strong and calm (Figure 3).

3. From both Figures 2 and 3, it may be seen that test misses and test hits are equally able to recognize overt expressions of abnormality.

---

[1]Adapted from a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, University of Minnesota, 1958.

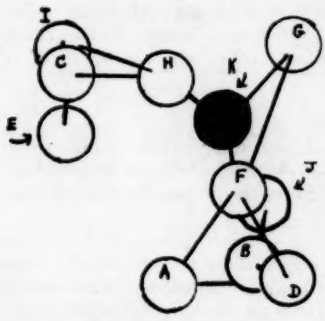


Fig. 1 : Normals

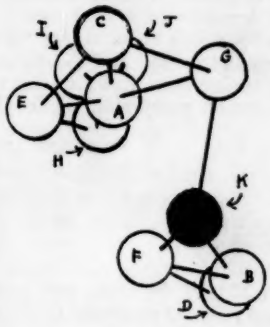


Fig. 2 : Test Hits

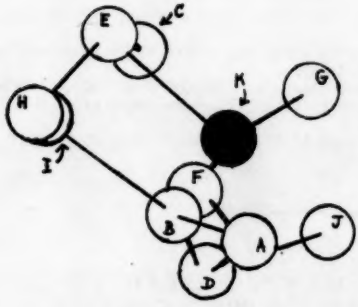


Fig. 3 : Test Misses

A- Me
B- They
C- A person who believes he is condemned
D- Friends
E- A person who is sure he gets a raw deal from life
F- People
G- Poison
H- People who would lie to get ahead
I- People who tell lies to keep out of trouble
J- The mind
K- The hypothetical concept which would be gotten by checking the midpoint position on all scales

Discussion

From the findings it seems evident that at least two distinct groups of paranoid schizophrenics exist and that one group considers itself normal and tends to resemble normals in its thinking about certain aspects of the world; whereas the other group describes itself as mentally ill. The former group corresponds to Guertin and Zilaitis'[1] Evasive and Well-Integrated Paranoid Type and the personality integration of this type is clearly illustrated in semantic differential performance. The test hits seem to correspond to the Grandiose and Delusional Paranoid Type found in the previous study and they are not reluctant to admit to mental illness.

Much speculation has been made by clinicians that the Evasive and Well-Integrated type is able to deny paranoid test items because he is able to recognize

these as overt expressions of abnormality. Although it has been demonstrated in this study that he does recognize such items, it has also been demonstrated that it is not recognition *per se* which accounts for the different MMPI *Pa* scale performances between the Well-Integrated and the Grandiose and Delusional Types. Both groups do seem to recognize these items as expressions of abnormality, but it is the Evasive and Well-Integrated group that denies these items.

Summary and Conclusions

1. Previous studies disclosed and cross-validated four types of paranoid schizophrenia. The current study identified at least two types and demonstrated that these groups, although apparently similar, think differently about themselves and environmental objects.

2. Semantic structures were plotted in three dimensional space and the differences between MMPI *Pa* scale test misses and test hits were noted. The former group described itself as mentally healthy, strong and calm, whereas the latter group described itself as mentally ill, weak and excitable. Both psychotic subgroups recognized overt expressions of abnormality and it was suggested that test misses choose to deny their symptoms whereas test hits admit to them.

3. On the basis of these findings it seems reasonable to suggest that the practice of bunching psychotics on the basis of symptoms is inadequate and does not assure the researcher a "pure" psychotic group.

References

1. Guertin, W. H. and Zilaitis, V. A transposed factor analysis of paranoid schizophrenics. *J. consult. Psychol.*, 1953, *17*, 455-458.
2. Guertin, W. H. A transposed analysis of the Bender-Gestalts of paranoid schizophrenics. *J. clin. Psychol.*, 1955, *11*, 73-76.
3. Osgood, C. E., and Suci, G. J. A measure of relation determined by both mean difference and profile information. *Psychol. Bull.*, 1952, *49*, 251-256.
4. Osgood, C. E. The nature and measurement of meaning. *Psychol. Bull.*, 1952, *49*, 192-237.
5. Osgood, C. E., Suci, G. J. and Tannenbaum, P. H. The *measurement of meaning*. Urbana: University of Illinois Press, 1957.
6. Wiener, D. N. Subtle and obvious keys for the MMPI. *J. consult. Psychol.*, 1948, *12*, 164-170.

BETA-WAIS COMPARISONS AND WAIS SUBTEST CONFIGURATIONS WITHIN A STATE PRISON POPULATION

JAMES H. PANTON

Reception Center, Central Prison
Raleigh, North Carolina

Problem

Large admission rates and understaffed classification sections force many prison institutions to rely on relatively short, group administered, and easily scored paper and pencil tests for appraisals of intelligence. One of the most widely used group intelligence tests in Federal and state penal institutions is the Revised Beta Examination developed for Army use during World War I, revised for civilian use by Kellogg and Morton[1] in 1934, and restandardized by Lindner and Gurvitz[2] in 1946. Wechsler's method of calculating IQ's is followed in the 1946 Beta restandardization by transforming raw scores into standard scores and then adding the standard scores with equal weights to obtain the total score. The total scores are then corrected to deviation IQ's with an SD of 15 and a mean of 100. The Beta being predominantly a non-language performance test raised the question as to its reliability in differentiating intelligence within a population where no additional IQ group test-

ing is conducted. The first phase of this study tests the reliability of the Beta by correlating it with Wechsler Adult Intelligence Scale (WAIS) results obtained on white and negro prison inmates. The second phase presents an analysis of WAIS protocols of white and negro prisoners in order to determine whether there are salient profile characteristics distinguishing personality patterns generally associated with indices of sociopathy.

Method

Samples of 100 white and 100 negro males were selected from the clinical files of the North Carolina Prison Department Reception Center in accordance with the distribution of Beta IQ's on 991 white and 769 negro male prisoners admitted to the Reception Center during the year 1958. All cases included in the samples had been administered the WAIS and Revised Beta Examination, and had been diagnosed as *Sociopathic Personality, Anti-Social Reaction* during their initial processing. Table 1 presents the comparison of the samples as to age, Beta IQ, education, and educational achievement as measured by the Wide Range Achievement Test (WRA). Inspection of the data reveals a close similarity between the samples and the 1958 admissions on all comparative factors. Longitudinal studies over a period of six years have shown that psychosociological characteristics of admissions to the Reception Center do not vary to any significant degree from year to year; therefore, it can be contended from the data presented in Table 1 that the samples are representative of the total prison population.

Table 1. Comparison of Means and Standard Deviations for Age, IQ, Education, and Educational Achievement Between White and Negro Experimental Samples and White and Negro 1958 Admissions

	Experimental Samples		1958 Admissions			
	Mean	S.D.	Mean	S.D.	M/Diff	*t*
Whites	(N = 100)		(N = 991)			
Age	28.15	10.64	27.88	10.37	.27	.24
Beta IQ	93.50	14.13	94.55	14.39	1.05	.70
Education	7.74	2.76	7.37	2.87	.37	1.27
Achievement (WRA)	5.84	1.23	5.49	2.61	.45	1.96
Negroes	(N = 100)		(N = 769)			
Age	28.20	11.07	28.43	9.94	.23	.20
Beta IQ	78.55	14.24	78.29	13.87	.26	.17
Education	6.46	3.30	6.82	3.18	.36	1.03
Achievement (WRA)	3.57	1.72	3.91	2.09	.34	1.80

The mean WAIS Verbal and Performance IQ's for each sample were compared with each other, and with the mean Beta IQ's. This procedure was followed by a comparison of the Full Scale WAIS and Beta mean IQ's for each sample. Pearson product moment correlations were computed for all comparisons. Mean scaled scores were computed for each subtest by sample, and the means were then ranked from highest to lowest value. Table 2 presents the WAIS-Beta mean comparisons, correlations between comparisons, and the rank order of the WAIS subtest scaled score means for each sample.

Results and Discussion

The data in Table 2 reveal that none of the mean differences appearing between WAIS-Beta IQ comparisons are statistically significant, and that with the exception of the negro Beta-WAIS Verbal correlation the Beta correlates at $r = .75$ and above with the WAIS Verbal, Performance, and Full Scale IQ's. The Beta-WAIS Full Scale correlations of .83 for the whites and .81 for the negroes imply that the Beta is a fairly reliable instrument for the measurement of IQ within prison populations where

TABLE 2. COMPARISON OF MEANS AND STANDARD DEVIATIONS AND CORRELATIONS BETWEEN VERBAL AND PERFORMANCE WAIS IQ'S AND BETA-WAIS FULL SCALE IQ'S, AND LISTING IN RANK ORDER THE WAIS SUBTEST SCALED SCORE MEANS WITH THEIR ACCOMPANYING STANDARD DEVIATIONS FOR WHITE AND NEGRO SAMPLES

Scales	A Mean	A S.D.	B Mean	B S.D.	M/Diff	t	r
			White Sample (N = 100)				
WAIS Verbal-Perf.*	90.75	14.02	91.60	14.14	.85	.42	.75
Beta-WAIS Verbal	93.50	14.13	90.75	14.02	2.75	1.37	.76
Beta-WAIS Perf.	93.50	14.13	91.60	14.14	1.90	.95	.84
Beta-WAIS Full Scale	93.50	14.13	90.55	13.61	2.95	1.49	.83
			Negro Sample (N = 100)				
WAIS Verbal-Perf.	77.20	10.63	78.55	14.21	1.35	.76	.75
Beta-WAIS Verbal	78.55	14.24	77.20	10.63	1.35	.76	.67
Beta-WAIS Perf.	78.55	14.24	78.55	14.21	—	—	.75
Beta-WAIS Full Scale	78.55	14.24	76.45	12.04	2.10	1.12	.81

Rank**	White Sample			Negro Sample		
Verbal	Subtest	Mean	S.D.	Subtest	Mean	S.D.
1	Comp	9.44	2.98	Dig Sp	7.14	2.45
2	Simil	8.76	2.81	Comp	6.67	2.13
3	Dig Sp	8.47	2.70	Simil	6.44	2.80
4	Arith	8.45	2.37	Arith	6.11	2.00
5	Info	7.58	2.20	Info	5.56	1.85
6	Vocab	7.54	2.67	Vocab	4.86	2.09
Performance						
1	Pic Comp	9.25	2.61	Obj Ass	7.28	2.71
2	Bl Des	9.12	2.45	Pic Comp	7.12	1.78
3	Pic Arr	8.89	2.71	Bl Des	7.02	2.60
4	Obj Ass	8.82	2.64	Pic Arr	6.65	2.24
5	Dig Sym	7.45	2.75	Dig Sym	5.45	2.94

*Scales on the left of the hyphen (-) are "A" column scales, and ones on the right are "B" column scales.

**In accordance with mean values.

limitations placed on the classification procedure make it necessary to employ easily administered group intelligence tests.

Wechsler[3] commenting on diagnostic groups, notes that sociopaths frequently score higher on WAIS Performance than on Verbal subtests; frequently make good scores on Picture Completion and score relatively low on Information, Similarities, and Digit Symbol. Although the WAIS sociopath characteristics noted by Wechsler are based primarily on the results from WAIS testing of adolescent sociopaths, it is Wechsler's contention that these same profile characteristics are applicable to adult sociopaths. The data of Table 2 reveal two characteristics opposing Wechsler's "sociopathic signs". These are (a) the relatively high ranking of the Similarities subtest means, and (b) the lack of statistical significance between the mean Verbal and Performance IQ's; however, in view of the fact that the mean often conceals more than it reveals it was decided to submit the subtest profiles to additional analysis before making any further interpretative deductions.

The two samples were next divided into subsamples according to whether the Verbal or Performance IQ exceeded the other to a significant degree. The table developed by Wolfensberger[4], and corrected by Fisher[5], for determining the significance of the difference between Wechsler Verbal and Performance IQ's was employed in separating the samples. This procedure produced a Verbal greater than Performance (V > P), a Performance greater than Verbal (P > V), and an indeterminate (V ≦ P) sample for both whites and negroes. The subtests were then ranked in accordance with the frequency in which each subtest scaled score appears either among the two highest or two lowest scaled scores. Table 3 presents the ranked sub-

tests for each of the subsamples. The percentage-in-rank figures for those subtests appearing consistently high or low throughout the subsamples are entered in the table. Whites and negroes present almost identical distributions within the subsamples. Only 37% of both samples present proposed sociopathic characteristics of Performance greater than Verbal, whereas 18% of the whites and 20% of the negroes display the reverse of this configuration. Over 40% of both samples are indeterminate in their Verbal-Performance relationship. These results would imply that the P > V indicator of sociopathy can only be relied upon to identify one out of three sociopaths and that three out of five will present profiles denoting no significant difference between Verbal and Performance IQ or will have Verbal IQ higher than Performance IQ.

The white and negro samples were next divided into subsamples in accordance with the highest 25%, middle 50%, and lowest 25% of the WAIS Full Scale IQ distribution. The same procedure for ranking subtest scaled scores as presented in Table 3 was employed in establishing the subtest configurations for each subsample presented in Table 4. Examination of the subtest rankings in tables 3 and 4 reveals that the white sociopaths present a high Comprehension, high Picture Completion, low Information, low Vocabulary, and low Digit Symbol profile regardless of the Verbal-Performance relationship or the level of mental development. The negroes present a profile consistently high on Comprehension and Digit Span, and low on Vocabulary, Information, and Digit Symbol. The high Comprehension-low Information and Vocabulary configuration appears unaffected by racial factors and holds true for all of the subsamples regardless of race with the exception of the Low 25%

TABLE 3. SUBTEST CONFIGURATIONS APPEARING AMONG SUBSAMPLES FORMED IN ACCORDANCE WITH VERBAL-PERFORMANCE IQ RELATIONSHIPS

Rank	Verbal > Perf. Subtest	% Rank	Perf. < Verbal Subtest	% Rank	Verbal ≦ Perf. Subtest	% Rank
		(N = 18)	White Sample (N = 37)			(N = 45)
Verbal						
1	Comp	100.0*	Comp	70.3	Comp	82.2
2	Dig Sp		Simil		Dig Sp	
3	Simil		Arith		Simil	
4	Arith		Dig Sp		Arith	
5	Vocab	55.6	Info	78.4	Vocab-	82.2
6	Info	72.2	Vocab	81.1	Info	84.4
Performance						
1	Pic Comp	72.2	Pic Comp	59.5	Pic Comp	71.1
2	Bl Des		Obj Ass		Bl Des	
3	Pic Arr		Pic Arr		Obj Ass	
4	Obj Ass		Bl Des		Pic Arr	
5	Dig Sym	61.1	Dig Sym	86.5	Dig Sym	82.2
		(N = 20)	Negro Sample (N = 37)			(N = 43)
Verbal						
1	Comp	100.0	Dig Sp	72.9	Comp	66.7
2	Dig Sp	72.2	Comp	62.2	Dig Sp	62.2
3	Arith		Airth		Simil	
4	Simil		Simil		Arith	
5	Info	77.8	Info	70.3	Info	66.7
6	Vocab	83.3	Vocab	94.6	Vocab	84.4
Performance						
1	Pic Arr		Obj Ass		Pic Arr	
2	Obj Ass		Bl Des		Pic Comp	
3	Pic Comp		Pic Comp		Bl Des	
4	Bl Des		Pic Arr		Obj Ass	
5	Dig Sym	83.3	Dig Sym	81.1	Dig Sym	80.0

*Percent of frequency in which the subtest score ranks either among the two highest or two lowest scores within each Verbal or Performance configuration.

IQ subsample of the negroes where Information is ranked among the two lowest scaled scores 40% of the time. The relatively high rank of Comprehension for both whites and negroes appears to support the often encountered ability of the sociopaths to rationalize their behavior.

It has been our experience that adult sociopaths will give those answers on the Comprehension items that they feel will be socially approved by the examiner. For example, in the envelope question (Item 3) 85% of the whites and 80% of the negroes give a two point response (mail it, put it in a mailbox, take it to the post office, etc.). Only 8% of both samples give zero point responses (open it, throw it away, leave it, etc.). The fire in the theatre question (Item 5) is responded to by a zero point answer (get the hell out of there, holler "fire" and run, etc.) by only 18% of the whites and 22% of the negroes. The whites answer Item 5 65% of the time with a two point response (inform the manager, tell the usher, etc.) and 17% of the time with a one point answer (turn in a fire alarm, try to put out the fire, etc.); whereas 41% of the negroes present two point answers to Item 5 and 39% present a one point response. The relatively low ranked Digit Symbol score of both samples seems to underscore the sociopathic characteristics of disdain for any task that requires persistent effort and concentration for any appreciable length of time. The low ranking of the Information and Vocabulary subtests on all but one of the subsamples lends support to Wechsler's contention that sociopaths are inclined to dis-

TABLE 4. SUBTEST CONFIGURATIONS APPEARING AMONG SUBSAMPLES FORMED IN ACCORDANCE WITH PERCENTAGE DISTRIBUTION OF WAIS FULL SCALE IQ'S.

Rank	Top 25% Subtest	Top 25% % Rank	Middle 50% Subtest	Middle 50% % Rank	Low 25% Subtest	Low 25% % Rank
			White Sample (N = 100)			
IQ Range	(101-122)		(82-100)		(50-81)	
Verbal						
1	Comp	84.0*	Comp	70.0	Comp	88.0
2	Arith		Simil		Dig Sp	
3	Simil		Arith		Arith	
4	Dig Sp		Dig Sp		Simil	
5	Vocab	52.0	Vocab	74.0	Info	72.0
6	Info	68.0	Info	88.0	Vocab	84.0
Performance						
1	Pic Comp	56.0	Pic Comp	66.0	Pic Comp	84.0
2	Pic Arr		Bl Des		Bl Des	
3	Obj Ass		Obj Ass		Obj Ass	
4	Bl Des		Pic Arr		Pic Arr	
5	Dig Sym	84.0	Dig Sym	86.0	Dig Sym	84.0
			Negro Sample (N = 100)			
IQ Range	(85-112)		(69-84)		(47-68)	
Verbal						
1	Dig Sp	64.0	Dig Sp	68.0	Comp	80.0
2	Comp	60.0	Comp	64.0	Dig Sp	72.0
3	Simil		Simil		Arith	
4	Arith		Arith		Info	40.0
5	Info	72.0	Info	78.0	Simil	
6	Vocab	88.0	Vocab	98.0	Vocab	68.0
Performance						
1	Bl Des		Pic Comp		Obj Ass	
2	Obj Ass		Bl Des		Pic Comp	
3	Pic Arr		Obj Ass		Pic Arr	
4	Pic Comp		Pic Arr		Bl Des	
5	Dig Sym	80.0	Dig Sym	78.0	Dig Sym	88.0

*Percent of frequency in which the subtest score ranks either among the two highest or two lowest scores within each Verbal or Performance configuration.

regard those aspects of educational and cultural training which do not lead to the satisfaction of immediate needs.

One of the most interesting aspects of the negro profiles is the consistently high rank of the Digit Span subtest in comparison with its fluctuation of position among the white subsamples. Seigman[6] has shown that Digit Span scores are affected by anxiety and Wechsler has also contended that low Digit Span scores, when not associated with organic defects, appear contributable in part to anxiety and inattention. Evaluation of the high ranking negro Digit Span scores in the light of these findings implies a relatively lesser degree of anxiety for the negroes than for the whites. This interpretation supports previous findings[7] at this institution, reporting negro inmates as demonstrating considerably less anxiety than white inmates. The whites present a high ranked Picture Completion score for all subsamples which appears to indicate a relatively greater ability to differentiate essential from non-essential details when performing a visual motor task than is demonstrated by the negroes.

Summary

The correlation between WAIS and Revised Beta Examination IQ's for white and negro samples of imprisoned sociopaths revealed that the Beta is a reliable predicator of IQ within an adult male prison population of low educational and cultural attainment. The WAIS sociopathic characteristic of Performance greater than Verbal IQ was found among only 37% of both white and negro prison samples, whereas 45% of the white sample and 43% of the negro sample presented WAIS profiles with neither Verbal nor Performance IQ's significantly higher than the other. 18% of the whites and 22% of the negroes presented profiles with Verbal greater than Performance.

An analysis of subtest scaled scores demonstrated that, irrespective of Verbal-Performance relationship ($P > V$, $V > P$, $V \leqq P$) or level of WAIS Full Scale IQ's, both whites and negroes frequently scored relatively high within their respective subsamples on Comprehension, and relatively low on Vocabulary, Information, and Digit Symbol. This configuration was interpreted as supporting the clinical impression of the sociopath in his attempts to receive immediate gain by "feeding back" to the examiner socially acceptable answers (high Comprehension); his disregard for educational and cultural training which does not lead to fulfillment of immediate needs (low Information and Vocabulary); and his difficulties in applying himself on tasks requiring persistent effort and concentration within an appreciable time frame (low Digit Symbol). The constantly high rank of Picture Completion of the whites was interpreted as implying a relatively greater ability than shown by the negroes to differentiate essential from non-essential details in performing visual motor tasks; whereas the consistently high rank of the Digit Span scaled score on the negro profiles was interpreted as implying a greater freedom from anxiety than was indicated in the white profiles.

References

1. Kellogg, C. E. and Morton, N. W. Revised Beta Examination. *Personnel J.*, 1934, *13*, 98-99.
2. Linder, R. M. and Gurvitz, M. Restandardization of the Revised Beta Examination to yield the Wechsler type IQ. *J. appl. Psychol.*, 1946, *30*, 649-65.
3. Wechsler, D. *The Measure and Appraisal of Adult Intelligence.* Fourth Edition. Baltimore: Williams and Wilkins Co., 1958.
4. Wolfensberger, W. P. Construction of a table for the significance of the difference between Verbal and Performance IQ's on the WAIS and the Wechsler-Bellevue. *J. clin. Psychol.*, 1958, *1*, 92.
5. Fisher, G. M., A corrected table for determining the significance of the difference between Verbal and Performance IQ's on the WAIS and the Wechsler-Bellevue. *J. clin. Psychol.*, 1960, *1*, 7-8.
6. Siegman, A. W. The effect of manifest anxiety on a concept formation task, a non-directed learning task, and a timed and untimed intelligence test. *J. consult. Psychol.*, 1956, *20*, 176-78.
7. Panton, J. H. Inmate personality differences related to recidivism, age, and race as measured by the MMPI. *J. correct. Psychol.*, 1959, *4*, 28-35.

THE VALIDITY OF WAIS PERFORMANCE SUBTESTS COMPLETED WITH ONE HAND[1]

PETER F. BRIGGS

University of Minnesota Medical School

Problem and Method

Cases with physical disabilities involving the upper extremities have an unknown handicap on tests requiring use of these affected limbs. Disabilities involving one arm are seen among stroke and polio patients with fair regularity. The range of such involvement varies greatly; for simplicity, this investigation simply considers the degree to which a patient is handicapped by having had available only the dominant or nondominant hand.

Procedures

Design. Four WAIS subtests require the use of the hands: digit symbol, picture arrangement, block design and object assembly. Actually, the digit symbol subtest requires but one hand for its performance if the test blank is held by the examiner, so the patient was tested on his dominant hand alone and on his nondominant hand alone. A counter balance design was developed whereby each subject was examined using his dominant hand alone, his nondominant hand alone, and both hands on the remaining three subtests. The possible combinations of the four tests under these conditions require twelve subjects for a complete block. There were four such blocks of male subjects and four such blocks of female subjects completed for this investigation. Thus, each of the 96 subjects took four subtests which netted a total of 384 subtest scores.

Subjects. The characteristics of 96 subjects collected at four different sources are given in Table 1. Twenty-four subjects, twelve males and twelve females, were tested from each source and the subjects from each source were kept separate within the analysis to permit the identification of variability due to the source. The first block of subjects were visitors to the Physical Rehabilitation Section of the University Hospitals, Minneapolis, Minnesota. There is an obvious difference between the

Table 1. Mean Age and Education of Groups Tested

Groups	Age		Education	
	Male	Female	Male	Female
Rehabilitation Visitors	42	30	11.9	11.2
Elizabeth Kenny and Masonic	45	42	9.6	10.4
Masonic	42	43	10.8	10.0
Staff	28	32	17.0	14.4
Total	39	37	12.3	11.5

mean age of the male (42 years) and female (30 years) subjects in this group. In education the sexes appear about equal. The second group of subjects was selected from patients at the Elizabeth Kenny Institute, Minneapolis, Minnesota and the Masonic Hospital at the University of Minnesota. The third set of subjects was from the Masonic Hospital and the fourth set of subjects was staff members at the University Hospitals and the Elizabeth Kenny Institute. Among this group the male sample is better educated than the female sample. This fourth set of subjects was included to test the importance of handedness at the upper time limit ranges on the performance tasks under consideration.

[1]This study was supported by a research grant from the Graduate School of the University of Minnesota.

In the selection of subjects, only volunteers were used with the examiner being able to elicit the cooperation of most of the patients and visitors approached. No volunteers were accepted with ages over 65 years or below 18. Each subject was asked before selection whether he had a hand preference and whether he had any disability involving either hand. Subjects with upper extremity disabilities or without hand preference were not selected.

Directions to the Subject. The subjects were told the purpose of the experiment was to find out how well they and other persons were able to perform certain tests with dominant and nondominant hands. They were asked which was their dominant hand and this hand was considered the dominant hand for the experiment. They were told that they would be given four subtests and that they would be informed as to which hand they could use for each of these. They were encouraged to keep the out-of-service hand in their lap during each subtest. After this preliminary briefing the standard subtest directions were administered along with each test in order printed on the answer sheet (2). A single examiner administered all the tests.[2]

Results

The scores subjected to analysis were weighted scores derived from each subtest after conversion by means of the standard tables on the front of the test blank. No attempt was made to correct these subtest scores for age. Therefore, age differences were expected within the data. Furthermore, as noted in the description of the subjects, the professional group was expected to be more able than the other subjects studied. The data were subjected to an analysis of variance in which the effect of the three hand conditions (dominant, nondominant, and both hands), sex, groups and subtests were identified along with the double and triple interaction for these effects.

Significant differences between groups and an interaction between sex and groups are shown in Table 2. The important considerations with respect to the major

Table 2. Analysis of Variance Effects with Significant F-Ratios

Source of Variation	df	Mean square	F	Probabliity
Hands	2	76.3	11.9	.01
Groups	3	140.3	64.1	.01
Tests	3	88.0	13.8	.01
Tests × Hands	5	23.5	3.7	.01
Sex × Groups	3	29.4	4.6	.01
Residual	296	6.4		

hypothesis were borne out in the significance of the hand conditions and of the interaction between subtests and hands. No interaction between groups and hands was noted. This is a failure to demonstrate that at the upper time limit ranges the absence of the more convenient hand is a handicap.

Table 3 presents the means of test scores for the three conditions of hand presentation indicating that when the nondominant hand is used in the digit symbol

Table 3. Means of Subtest and Hand Combinations

Subtest	Hand Combination			Total
	Dominant	Nondominant	Both	
Digit Symbol	9.4	6.2		7.8
Block Design	10.3	9.9	9.8	9.9
Picture Arrangement	8.7	8.4	9.0	8.7
Object Assembly	9.2	9.6	10.2	9.7
Total	9.4	8.2	9.7	

[2]The author wishes to thank Mr. David Speer who did the testing for this study.

subtest, an average penalty of over three scale score points results. No other subtest seems to be affected by the single handed performance or nondominant hand performance to a significant degree. There is a slight effect on the object assembly subtest and yet the means do not suggest that the deflection from the both-handed performance is meaningful since the dominant hand performance is poorer than that of the nondominant hand. It is obvious both from the analysis of variance and from inspection of the means that the picture arrangement subtest is more difficult than the other subtests for this group as already noted by Garfield.[1]

Discussion

The availability of only one hand when this hand is the dominant hand does not seem to affect the performance subscales on the WAIS. When only the nondominant hand is available, three of the four performance subtests involving the use of the hands appear to be valid. Only the digit symbol subtest clearly requires the use of the dominant hand for its valid administration. It is generally rewarding to find that the aspects of intelligence sampled by this particular test are so little dependent on hand coordination. There is no consideration of the reason for hand loss, since all the subjects in this experiment were normal both with respect to intellectual function and hand function.

Summary

This study compared the importance of having a dominant or a nondominant hand available as contrasted with having both hands available for the completion of the four WAIS subtests: digit symbol, object assembly, block design, and picture arrangement. Ninety-six subjects were examined on these four subtests in a counterbalanced incomplete block design. The dominant, nondominant, and both hands were sampled. The subjects for the most part were hospital patients or hospital visitors with a sample of high IQ individuals incorporated to test the importance of these results at extreme ranges.

The subtest scale score results were subjected to an analysis of variance. It was found that the digit symbol subtest was greatly affected by the handicap of using the nondominant hand; the subject was penalized by an average of 3 points. No other subtest showed sensitivity to handedness when comparing the dominant hand with the nondominant hand or comparing a single hand performance with a two hand performance. The picture arrangement subtest was found to be more difficult than the other three subtests for the group as a whole. This result was not interpreted. The experimenter concluded that handicapped subjects can be given the full WAIS if they have their dominant hand intact and all parts of the WAIS except the digit symbol test if they are working with the nondominant hand.

Bibliography

1. Garfield, Sol L. An appraisal of object assembly on the Wechsler-Bellevue and the WAIS. *J. clin. Psychol.* 1960, *16*, 8-9.
2. Wechsler, D. *Manual for the Wechsler Adult Intelligence Test.* New York: Psychological Corporation, 1955.

THE PERFORMANCE-VERBAL IQ DISCREPANCY IN A GROUP OF SOCIOPATHS

J. G. FIELD[1]

Institute of Psychiatry, Maudsley Hospital
University of London

Problem

Wiens, Matarazzo and Gaver [9] observe that "many clinicians assume that the intelligence tests of sociopaths[2] will be characterized by a Performance IQ that exceeds the Verbal IQ." They quote five studies on the Wechsler Bellevue test [8] of which one [1] supports this assumption and four do not [2, 5, 6, 7]. Wiens *et al.* then go on themselves to test the assumption that sociopaths earn a significantly higher performance IQ than verbal IQ. Their results are positive, and they conclude that when the criterion of sociopathy is clearly defined by overt behavior the assumption will be supported. The criterion of sociopathy used by them was that each subject had been convicted in a court of a sexual offence against another person.

Method

In an attempt to repeat Wiens's finding, the Wechsler-Bellevue IQs of men serving sentences of Preventive Detention were examined.[3] These men had been consecutively received at a prison during a twelve month period. In the United Kingdom a person may be sentenced to preventive detention if he is not less than thirty years of age, is convicted of an offense punishable with imprisonment for at least two years, and has been convicted on at least three previous occasions after the age of 17 of offenses likewise punishable, and was on at least two of these occasions sentenced to certain forms of detention. The court must also be satisfied that it is expedient for the protection of the public that the person should be detained for a substantial time. A sentence of preventive detention may not be of less than five or more than fourteen years [3].

All men whose native tongue was not English, or who had had previous experience of the Wechsler-Bellevue were excluded. All men who had no court appearances before their twenty-first birthday, or who had fewer than ten previous convictions, were likewise excluded. This yielded sixty-five subjects (Mdn age = 39.5), a group which provides sufficient degrees of freedom to repeat Wiens' finding at the same level of significance. Such criteria of sociopathy are, if anything, an improvement on those used by Wiens, in the sense of being more extreme.

Results

Table 1 gives our results with Wiens' results given in brackets for comparison. In the present sample the verbal and performance IQs are identical to all significant places. Wiens' observation that a verbal-performance IQ discrepancy will be found if behavioral criteria of sociopathy are used, is not supported.

It can be noted that the present sample has higher than average IQs. In the United Kingdom it is not the practice to sentence mental defectives to prison, and persons bordering on mental deficiency may therefore be under-represented in our sample. This explanation would also account for the attenuation at the lower end of the IQ ranges in the present sample and the smaller standard deviations. In Wiens' sample 7% were diagnosed mental defectives.

---

[1]The author gratefully acknowledges the cooperation of the Prison Commission, and the prison governors and psychologists. This investigation is part of a project supported by the Medical Research Council.

[2]Sociopath = "a broad category for disorders in one's relationship with society and with the cultural milieu" [8].

[3]The testing was carried out by the prison psychological staff, and the author gratefully acknowledges permission to use the results.

TABLE 1. PERFORMANCE-VERBAL IQ DISCREPANCIES*

	Performance IQ	Verbal IQ	Total IQ	t	P
Mean	104.1 (104.0)	104.0 (97.6)	104.7 (101.1)	N. S. (5.41)	(.001)
S. D.	12..1 (15..0)	14..6 (17..3)	13.1 (15.3)		
Range	80 - 130 (62 - 137)	75 - 141 (69 - 130)	77 - 134 (64 - 135)		
% Ss	P>V 46 (72)	V>P 52 (28)		N. S. (-)	(..001)

*The results of Wien *et al.* are given in parentheses.

SUMMARY

Wiens *et al.* report a significant verbal-performance IQ discrepancy on the Wechsler-Bellevue scale for sociopaths selected on a behavioral criterion. The present study fails to repeat the finding on a group selected on more extreme criteria of sociopathy.

REFERENCES

1. CLARK, J. H. Subtest variation on the Wechsler-Bellevue for two institutionalized behavior problem groups. *Amer. Psychologist*, 1949, *4*, 395, (Abstract).
2. CLARK, J. H. and MOORE, J. H. The relationship of Wechsler-Bellevue patterns to psychiatric diagnoses of Army and Air Force prisoners. *J. consult. Psychol.*, 1950, *14*, 493-495.
3. Criminal Justice Act, Section 21, (2), 1948. Her Majesty's Stationery Office, London.
4. ENGLISH, H. B. and ENGLISH, A. C. *A Comprehensive Dictionary of Psychological and Psychoanalytical Terms.* New York: Longmans Green, 1958.
5. GURVITZ, M. S. The Wechsler-Bellevue test and the diagnosis of psychopathic personality. *J. clin. Psychol.*, 1950, *6*, 397-401.
6. STROTHER, C. R. The performance of psychopaths on the Wechsler-Bellevue test. *Proc. Ia. Acad. Sci.*, 1944, *51*, 397-400.
7. THURSTON, J. R. and CLADEN, G. Intelligence factors in irregular discharge among tuberculosis patients. *J. consult. Psychol.*, 1954, *18*, 404.
8. WECHSLER, D. *The Measurement of Adult Intelligence* (3rd Ed.). Baltimore: William & Wilkins, 1944.
9. WIENS, A. N., MATARAZZO, J. D. and GAVER, K. D. Performance and verbal IQ in a group of sociopaths. *J. clin. Psychol.*, 1959, *15*, 191-193.

THE RELATIONSHIP BETWEEN THE CMAS AND WISC FUNCTIONING[1]

A. JACK HAFNER — *University of Minnesota*

DONALD M. POLLIE — *Arizona State University*

IRWIN WAPNER — *Ritenour Consolidated School District*

PROBLEM

Two previous studies by McCandless and Castaneda[(2)] and Hafner and Kaplan[(1)] have reported significant negative correlations between the Children's Manifest Anxiety Scale (CMAS) and intellectual functioning as measured by the Otis. The present investigation was concerned with the relationship between an individually administered intelligence test and CMAS scores. The test chosen was the Wechsler Intelligence Scale For Children (WISC) as it permitted a more detailed evaluation of the effects of manifest anxiety on various kinds of intellectual functioning.

PROCEDURE

The WISC was administered to 42 children, nine girls and 33 boys. Their ages ranged from 10-12 years with a mean age of 10.8 years. Twenty-four of the children were drawn from a child guidance clinic and the other 18 children were tested in

[1]The data for this study were collected while the authors were in St. Louis, Missouri.

public schools as part of a screening for an enrichment program or because of scholastic difficulties. The CMAS was administered at the same time of the WISC testing to all the subjects.

Results and Discussion

Table 1 shows the test results in terms of means and standard deviations for the group. The product-moment correlations between the CMAS and the WISC

Table 1. CMAS and WISC Means and SDs

	Mean	SD
Anxiety Scale	16.81	7.14
L Scale	3.02	2.41
WISC	105.92	17.34

scales and subtests ranged from +.002 for Picture Completion to −.34 for Coding. Only the correlations for Block Design (−.31) and Coding (−.34) were significant at the .05 level.

Table 2. Product-Moment Correlations Between CMAS Scores and WISC Functioning

Item	r
Full Scale IQ	−.26
Verbal Scale IQ	−.21
Performance Scale IQ	−.28
Information	−.20
Comprehension	−.17
Arithmetic	−.16
Similarities	−.22
Vocabulary	−.23
Digit Span	−.12
Picture Completion	+.002
Picture Arrangement	−.19
Block Design	−.31*
Object Assembly	−.20
Coding	−.34*

*Significant at .05 level

While the correlation of −.26 between the CMAS and the WISC Full Scale IQ did not quite reach significance at the .05 level, the magnitude of the r is similar to the significant r of −.21 found by Hafner and Kaplan[1] using a larger population and the Otis. Except for the Picture Completion subtest, the consistent negative correlations between the WISC items and the CMAS suggest a general negative relationship between manifest anxiety and WISC functioning.

It is interesting to note that the two significant correlations between WISC subtests and the CMAS were for Block Design and Coding. These are the two Performance Scale subtests which are most frequently interpreted clinically as being susceptible to the influence of anxiety. The two Verbal Scale subtests of arithmetic and digit span which are also frequently interpreted in this manner, however, did not yield significant correlations with the CMAS.

Summary

The CMAS and WISC were administered individually to 42 children with a mean age of 10.8 years. The correlations between the CMAS and WISC scales and subtests were all negative except for the Picture Completion subtest. Only the correlations between the CMAS and the subtests of Block Design and Coding reached significance at the .05 level.

References

1. Hafner, A. J. and Kaplan, A. M. Children's manifest anxiety and intelligence. *Child Develpm.*, 1959, *30*, 269-271.
2. McCandless, B. R. and Castaneda, A. Anxiety in children, school achievement, and intelligence. *Child Develpm.*, 1956, *27*, 379-382.

RORSCHACH INDICES AND AUTONOMIC STRESS REACTIVITY*

J. E. TONG AND I. C. MURPHY

Rampton Hospital, Near Retford, Notts., England

INTRODUCTION

Controversy over the validity and utility of the Rorschach test in clinical work is likely to remain current for some time to come. Complete rejection of the test may yet be premature [3, 4], and in assessing the results of validation studies one should bear in mind that several possibilities exist. Among these is the possibility that the test may have some value only when applied to certain nosological groups, or for the purpose of isolating particular characteristics. Secondly, the usual validation studies may be insufficiently acute in that they frequently confuse the ability of the Rorschach test to isolate certain personality variables and the hypothetical relationship of these variables to the clinical diagnostic or prognostic problem in hand. There have been few studies aimed to assess the ability of the Rorschach test to isolate particular variables without consideration of the relationship of these variables to clinical problems, and there have still been relatively few attempts to validate the numerous Rorschach indices each in turn [2].

This paper presents the results of a correlation study comparing Rorschach indices of anxiety and emotional activity with different forms of response to different levels of stressors. A similar study by Lacey [10] indicated considerable support for the Rorschach claim to demonstrate emotionality by manipulation of the color score, for normals. In the problems of delinquent and psychopathic behavior, laboratory stress reactivity and general emotional responsiveness have been shown to be closely related to both diagnostic and prognostic issues [6, 11, 12, 13, 14] and it is of some practical importance to establish the extent to which the Rorschach indices are related to physiologically defined variables.

PROCEDURE

Subjects. These are 200 male patients at a State Hospital for mental defectives of dangerous or violent propensities; age range 15-50 years; modal age 22; mean IQ 76, psychotic and organic cases strictly excluded.

The men had all been responsible for behavior proscribed by criminal law, usually serious and repetitive and many of the men could be described as psychopathic. Each patient was completely free to participate in any examination or not, and all remained contented and cooperative as is necessary for valid results with the particular techniques, which are, of course, essentially innocuous.

Techniques. All Ss were examined routinely with the Rorschach test using standard scoring methods [8, 9] and also by one or more of several laboratory tests. All testing for any one patient was completed in a day, and the various measures assessed independently on a later occasion. The laboratory indices to be considered are:

1. Autonomic response to a generalised stressor: skin temperature change during an attempt to solve an insoluble problem [14]. The thermometer probe is placed on the back of the left wrist under a loose rubber band. The subsequent data also refer to the actual temperature reading immediately prior to commencement of the insoluble task.

2. Autonomic response to a specific stressor: the number of conditioned and unconditioned galvanic skin responses during a standard conditioning session using auditory stimulus for the C. S. and slight touch to the closed eye with cotton wool as the U. C. S. [14].

3. Leg persistence time: a fatigue-pain resistance test measuring the length of time the patient holds his extended right leg above the seat of a chair whilst seated in a chair of similar height [13].

The Rorschach indices to be considered are: (a) the total number of responses to the cards; (b) the total number of achromatic responses, i.e. FC^1, C^1F, C^1; (c) the total number of shading

*Supported by a Medical Research Council Scholarship.

responses, i.e. Fc, cF, c, FK, KF, K, Fk, kF, k; (d) the total number of color responses, i.e. FC, CF, C; (e) the score "Sum C", which accords the value of 0.5, 1.0, and 1.5 for each FC, CF, and C response respectively; (f) the total reaction time for the chromatic cards, assessed only for those S's who responded to all the cards; (g) the total reaction time for the achromatic cards, assessed for those S's who responded to all the cards; (h) reaction time (chromatic) minus reaction time (achromatic), i.e. an index of "color shock"; (i) the total number of achromatic cards not eliciting a response; (j) the total number of chromatic cards not eliciting a response.

Correlation analyses were undertaken for each laboratory index against each Rorschach index by means of the product-moment method, Kendall's tau, or chi square. In the table giving the results where a significant value for chi square is quoted this is based upon a standard tetrachoric form for the variables with each relevant variable being dichotomized as follows: (a) temperature change classified as "decrease greater than 0.1° C" or "no change or increase", (b) the number of Rorschach responses classified as "< 15" or "> 14"; (c) the number of achromatic responses, shading responses, chromatic responses, rejections of achromatic cards and rejections of chromatic cards classified as "zero" or "greater than zero"; and (d) the "Sum C" score classified as "< 1" or "> 0.5". In the same table, where chi square is stated to be not significant, various methods of treating the raw data were adopted none of which produced a sufficiently high statistic.

Results

The results are summarized in Table 1, and two points are important in consideration of these data. First, when a fairly large number of analyses is undertaken one can expect that a few co-efficients might reach the 5% level of significance, on chance grounds [3, 4]. This consideration is of less importance when some of the results show a fairly consistent trend, as in this case. In the second place, all Rorschach variables, with the exception of the temporal factors, are likely to be correlated with the general level of responsiveness indicated by the total number of R's.

With these points in mind, the following features emerge from the data. (a) In general there is very little or no relationship between Rorschach indices of anxiety and emotionality and laboratory indices of stress reactivity with the exception of: (i) shading and achromatic responses which are slightly associated with the laboratory measures, and, (ii) color responses which are highly associated with laboratory measures. (b) The only laboratory index which is consistently related to Rorschach indices is the skin temperature reaction to experimental frustration.

There is a high correlation ($\chi^2 = 5.89$; $P = .02$) between the total number of Rorschach responses and the skin temperature change, (*i.e.*, temperature fall (the stress response) is associated with relatively high Rorschach responsiveness), in addition to the somewhat higher correlation ($\chi^2 = 11.46$; $P = .001$) between the "Sum C" score and the temperature change. However, there is also a high positive association between general Rorschach responsiveness and "Sum C", in this particular group, ($\chi^2 = 7.34$; $P = .01$). We feel that this latter correlation does not necessarily suggest a causal connection and may indicate association with a common but more fundamental personality variable. Nevertheless, it is important to show whether or not there is a relationship between the skin temperature reaction and "Sum C", independent of general Rorschach responsiveness. The raw data do not lend themselves to partial correlation methods, so that it was necessary to restrict the analysis to those subjects giving a total number of responses between 11 and 20, inclusive. These result in the following contingency table for which the value for chi square is 4.83, $P = .05$, and indicates support for the Rorschach hypotheses linking the "Sum C" score with emotional responsiveness.

	Sum C	
	< 1.0	>0.5
Temp. decrease > .1°C	3	15
Temp. increase or no change	8	5

TABLE 1. CORRELATIONS BETWEEB RORSCHACH INDICES AND LABORATORY TEST SCORES

Rorschach Index	Frustration Reactions: Basal Temperature				Frustration Reactions: Temperature Change [a]				G. S. R. Conditioning: Σ R				G. S. R. Conditioning: Σ C. R.				Leg Persistence Time			
	N	Stat.	Value	P [b]	N	Stat.	Value	P	N	Stat.	Value	P	N	Stat.	Value	P	N	Stat.	Value	P.
Total Number of Responses	71 71	r tau	−0.10 −0.07	NS NS	68	X^2	*** 5.89	.02	171 171	r tau	+0.07 +0.01	NS NS	171 171	r tau	+0.08 +0.06	NS NS	154 154	r tau	+0.11 +0.07	NS NS t
Number of Achromatic Responses	71	X^2	*	NS	72	X^2	*** 3.70	.10	172	X^2	** 4.79	.05	173	X^2	*	NS	155	X^2	*	NS
Number of Shading Responses	72	X^2	*	NS	72	X^2	*** 4.37	.05	172	X^2	** 5.62	.02	174	X^2	** 3.10	.10	156	X^2	*	NS
Number of Color Responses	72	X^2	5.11	.05	73	X^2	*** 4.47	.05	172	X^2	*	NS	172	X^2	*	NS	158	X^2	*	NS
Sum C	71	X^2	*	NS	70	X^2	*** 11.46	.001	171	X^2	*	NS	171	X^2	*	NS	156	X^2	*	NS
Color Shock R.T. (C) - R.T. (A)	41 41	r tau	−0.14 −0.03	NS NS	41	X^2	*	NS	86 86	r tau	+0.11 +0.03	NS NS	86 86	r tau	−0.04 −0.05	NS NS	84 84	r tau	−0.08 −0.12	NS NS
Reaction Time—Achromatic	41 41	r tau	−0.21 −0.10	NS NS	41	X^2	*	NS	86 86	r tau	−0.01 −0.02	NS NS	86 86	r tau	+0.09 +0.03	NS NS	84 84	r tau	+0.02 +0.05	NS NS
Reaction Time—Chromatic	41 41	r tau	−0.08 −0.09	NS NS	41	X^2	*	NS	86 86	r tau	+0.01 −0.05	NS NS	86 86	r tau	+0.06 +0.03	NS NS	84 84	r tau	+0.05 +0.03	NS NS
Rejections—Achromatic	73	X^2	** 5.85	.02	72	X^2	** 3.92	.05	172	X^2	*	NS	172	X^2	*	NS	158	X^2	*	NS
Rejections—Chromatic	71	X^2	*	NS	72	X^2	*	NS	172	X^2	*	NS	172	X^2	*	NS	158	X^2	*	NS

Key [a]Dichotomized into "Decrease > .10C" and "No change or Increase". A positive association indicates an association between Temperature Decrease and a low Rorschach score, and conversely for a negative association.

[b]Two-tailed test.

*Several values were obtained by dichotomizing the variable at different points, but no value was significant.

**Positive association. ***Negative association. NS = Not significant.

Discussion

The results justify the current opinion that many Rorschach indices are not valid indicators of emotional reactivity, anxiety, etc., but nevertheless suggest that the test contains some features worth developing in a more sophisticated and rigorous manner. At this stage it may be advisable not to assume that the high positive correlation between general responsiveness and "Sum C" indicates a *causal* connection, and assume that there are two factors in operation. Similar evidence of a specific color factor has also been reported for normals[10] and a recent review of the literature[1] quotes no study indicating a positive correlation between the total number of responses and color. With psychopaths one would expect low Rorschach responsiveness to be shown by those who are unreactive to laboratory stressors, these also being the passively uncooperative type least responsive to therapy and training and analogous to primary sociopaths[11, 14]. A learning theory explanation of why those prone to temperature fall during frustration should also be responsive to the color of the cards may be found in the obviously early and well-established conditioned link between the perception of blood and pain. Neurotic reaction to blood is discussed by Haward[6] in relation to Rorschach responses, and the whole concept of color responsiveness has been well examined by Keehn[7].

The second feature to emerge concerns the indication that the significant correlations predominantly exist for the Rorschach indices and the skin temperature index, and not for the other laboratory stress measures. This is of both practical and theoretical interest. For practical purposes this could explain why the Rorschach is found to be a poor prediction tool[3, 4]. The test is clearly measuring a particular anxiety factor, similar to that of skin temperature reaction to frustration which is known to correlate less highly with social criteria than measures such as G. S. R. conditioning, pain perception and pain resistance[13, 14]. Speculation as to the nature of this particular anxiety factor leads us to suggest that it differs from the other stressor response situations by representing higher order conditioning. If we regard anxiety as the conditioned fear response to pain as the primary stimulus, which in turn functions as a stimulus then, by definition, more pure measures of anxiety will be isolated by techniques which make the stimulus fairly specifically or closely linked to the definition, and apply measurement at the autonomic or restricted overt level. On the other hand reaction to an insoluble problem in the clinical situation is reaction to an ego involvement situation which is well removed from, and involves drives not present in, the basic definition. Laboratory reaction to this sort of situation need not be indicative of social or clinical behavior except that which is a reaction to ego involvement situations. It should be noted, however, that in the only comparable study[10] the stressor response index found to correlate with Rorschach color was not an ego involvement situation, but the cold pressor test.

Summary

Two hundred male psychopaths certified as mentally defective were examined with the Rorschach test and with laboratory stress techniques, and the scores subjected to correlation analyses for the purpose of validating Rorschach indices of anxiety and emotional reactivity. Some slight support was found for the hypothesis relating Rorschach shading responses to anxiety level, and greater support for the hypothesis relating the color scores to stress reactivity. The correlations are seen to be suggestive of a particular anxiety reaction associated with ego involvement, and in general the data are contra-indicative of a high validity for the Rorschach test in its present form.

References

1. Baughman, E. E. The role of the stimulus in Rorschach responses. *Psychol. Bull.*, 1958, *55*, 121-147.
2. Cronbach, L. J. *Essentials of Psychological Testing*. New York: Harper, 1949.
3. Eysenck, H. J. Learning theory and behaviour theory. *J. Ment. Sci.*, 1959, *105*, 61-75.
4. Eysenck, H. J. Personality Tests: 1950-1955. In *Recent Progress in Psychiatry*. G. W. T. H. Fleming and A. Walk (Eds.). London: Churchill, 1959.

5. HAWARD, L. R. C. Color associations as a contributing factor to neurotic color shock. *Brit. J. Med. Psychol.*, 1955, *28*, 183-187.
6. KAY, B. Reactions of delinquents and other groups to experimentally produced frustration. *Brit. J. Delinq.*, 1954, *4*, 245-253.
7. KEEHN, J. D. Rorschach validation III: an examination of the role of color as a determinant in the Rorschach test. *J. Ment. Sci.*, 1953, *99*, 410-438.
8. KLOPFER, B. and KELLEY, D. Mc. G. *The Rorschach Technique*, New York. World Book Co., 1946.
9. KLOPFER, B., AINSWORTH, M. D., KLOPFER, W. G. and HOLT, R. R. *Developments in the Rorschach Technique*, Vol. 1. London, Toronto, Wellington, Sydney: Harrap, 1954.
10. LACEY, J. I., BATEMAN, D. E. and VAN LEHN, R. Autonomic response specificity and Rorschach color responses. *Psychosom. Med.* 1952, *14*, 256-260.
11. LYKKEN, D. T. A study of anxiety in the sociopathic personality. *J. abnorm. soc. Psychol.* 1957, *55*, 6-10.
12. TONG, J. E. Perceptual (Rorschach) characteristics of the sexual offender. *Brit. J. Delinq.* 1955, *4*, 304-307.
13. TONG, J. E. The assessment of anxiety as an intervening variable in the delinquent behaviour of M. D. subjects. Galvanic skin response and leg-persistence indices. *Brit. J. Psychol.*, 1957, *48*, 13-25.
14. TONG, J. E. Stress reactivity in relation to delinquent and psychopathic behaviour. *J. Ment Sci.*, 1959, *105*, 935-956.

THE PROJECTIVE EFFECTS OF FEAR AND SEXUAL AROUSAL ON THE RATING OF PICTURES[1]

ROBERT H. HORNBERGER[2]

State University of Iowa

While the concept of projection is very widely used in psychology and related disciplines, it has been variously defined. Of particular interest to this study is the distinction between similar and complementary projection[3]. In "similar projection" a person attributes his characteristics to others while in "complementary projection" a person attributes to others such traits as would justify or necessitate his typical reaction to them. According to similar projection, then, a frightened person would perceive others as "frightened," but if complementary projection were operating he would see them as "frightening." Most relevant to the present study is the classic work of Murray [5] in which he found evidence for the complementary projection of fear.

In spite of the various definitions of projection, the process of projection is generally agreed to be a function of both the characteristics of the projecting organism and of the instigating stimulus. One stimulus property which has received considerable attention is that of stimulus ambiguity; however, the results of the studies [3, 6] of this factor have been equivocal.

This study was designed to provide additional information about the process of projection. Specifically an attempt was made to explore the usefulness of the distinction between similar and complementary projection and to gather additional information about the effects of stimulus characteristics in projective responses.

PROCEDURE

Seventy-two male college students ranging in age from 18 to 27 years (median age = 21.0) were divided into three experimental groups of 24 *Ss* each. All *Ss* were run individually.

---

[1]This paper was presented at the Rocky Mountain Psychological Association meetings in Santa Fe, May, 1958, and is based on a doctoral dissertation [12] done at the State University of Iowa. Thanks are due particularly to Leonard D. Goodstein for his assistance during the study.

[2]Now at the University of Nebraska College of Medicine.

[3]Campbell, D. T. Five varieties of projection in trait attribution. USPHS, Project M-1544. Personal Communication. April, 1958.

The Fear Arousal group (N = 24) was threatened by complex electrical equipment, the instructions, strapping the *S* in a chair designed to look like an electric chair, placing electrodes on his head and one hand, shocking him on the fingers during a preliminary task, and threatening him with additional shock during the picture rating procedure. However, no shock was actually administered while the *S* was rating the neutral pictures as described below. The Sexual Arousal group (N = 24) was exposed to an attractive woman who dressed and behaved in a provocative manner and who had each *S* rate pictures of nude women for beauty immediately prior to the administration of the neutral pictures. The Control group (N = 24) rated the neutral pictures without prior experimental activity.

All *S*s rated portraits of eight male and eight female college students which were projected individually onto a screen. These pictures were selected for their neutrality or ambiguity of facial expression. The *S* selected ten adjectives to describe each picture, one adjective from each of ten groups of three. Three of these triads contained a fear similar, a fear complementary, and a buffer adjective; three of them contained a sex similar, a sex complementary[4], and a buffer adjective; and four of the triads were buffer triads, i.e., contained no fear or sex adjectives. The ambiguity of the pictures was varied by exposing them at time intervals of either .2 or 5 seconds. Two sets of four male and four female pictures (Set I and Set II) were formed and each of the three experimental groups was divided into two sub-groups. The exposure time was then counterbalanced with the sets of pictures.

Results

Interviews with the *S*s of the fear arousal and sexual arousal groups indicated that they were not aware of the purpose of the experiment. Furthermore the fear arousal procedures appeared to be quite effective in frightening the *S*s, but only half of the *S*s in the sexual arousal group manifested some degree of sexual arousal. Many of the *S*s in the latter group seemed surprised and /or embarrassed over the procedures. The statistical analyses performed on the data were extensions of a Lindquist (4) Type VII analysis of variance.

Table 1. Mean Experimental Treatments Effects for Three Groups*

Score	Fear Arousal	Sexual Arousal	Control	*P*
Fear Similar	3.55	3.55	4.18	NS
Fear Complementary	3.83	3.36	3.05	$<.05$
Sex Similar	3.70	3.10	3.67	NS
Sex Complementary	4.01	4.09	4.38	NS

*The *M*s in this table are based on the complete data, collapsed over the other factors. The basic unit in the statistical analyses was the total number of responses of a particular kind, fear similar, etc., made to the four similarly sexed pictures of a given set.

Table 1 shows that the fear arousal, sexual arousal, and control groups differed significantly among themselves only in the fear complementary scores. The fear arousal group rated the neutral pictures as "dangerous", "frightening", and /or "threatening" more frequently than either the sexual arousal or control groups although only the difference between the fear arousal and control groups was significant ($p < .05$). While no significant differences were found among the experimental groups for the sex complementary scores, there was a tendency for the sexual arousal group to give fewer sex similar responses than either the fear arousal or the control group. That is, the *S*s of the sexual arousal group rated the neutral pictures as

[4]"Afraid", "frightened", and "scared" were the fear similar adjectives; "dangerous," "frightening," and "threatening," the fear complementary adjectives; "aroused," "desirous," and "passionate," the sex similar adjectives; and "appealing," "arousing," and "exciting," the sex complementary adjectives.

"aroused", "desirous", and/or "passionate" less often than the *Ss* in the fear arousal or control groups.

For all four scores the pictures of Set I were responded to differently from the pictures of Set II; however, the only significant effect of the exposure time of the pictures was that the pictures were rated as "appealing", "arousing", and/or "exciting" more often under the short exposure time than under the long exposure time. All of these differences were statistically significant at the 5% level or less.

The female pictures were rated as "afraid", "frightened", and/or "scared" more often than the male pictures. Conversely, the men were rated as "dangerous", "frightening", and/or "threatening" more frequently than the women and also the women were rated as "aroused", "desirous", and/or "passionate" more often than the men. All of these sex differences were statistically significant at less than the 1% level.

Discussion

That the fear arousal group rated the pictures as "dangerous", "frightening", and/or "threatening" more frequently than the sexual arousal and control groups substantiates Murray's[5] findings and constitutes additional evidence for the complementary projection of fear. There is, however, no evidence for the complementary projection of sexual arousal or the similar projection of either motive. In fact, the findings that the fear arousal group made fewer fear similar responses than the control group, and that the sexual arousal group obtained lower sex similar scores than the control group suggest an inhibition of similar projection. That the only evidence of projection was of the complementary projection of fear suggests that the occurrence and/or type of projection may be related to the motivation involved. The present study indicates the need to distinguish between different types of projection and to investigate the conditions under which each occurs.

That the *Ss* of the sexual arousal group made fewer sex similar responses than the *Ss* of either the fear arousal or control groups indicates an inhibition of sex similar responses by the *Ss* of the sexual arousal group and is probably related to the surprise and/or mild embarrassment over the experimental procedures which some of the *Ss* of this group manifested in the interview. This inhibitory effect is similar to Clark's[1] finding that the *Ss* of his sexual arousal group related significantly fewer sex and guilt themes under non-alcoholic conditions than did the *Ss* of his control group. However, the present data provide no ready answers to the questions of why the occurrence of complementary projection was limited to the *Ss* of the fear arousal group and why similar projection occurred in neither group.

The finding that the pictures of Set I were rated significantly differently from the pictures of Set II for all of the four scores indicates that the two sets of pictures varied systematically in their value as stimuli for the different responses investigated in this study. Also the differences in the ratings of the male and female pictures further emphasize the importance of stimulus characteristics; these sex differences can be understood in terms of the different roles assigned to men and women in American culture. However, the failure of the exposure time to be a significant factor in any of the scores but the sex complementary one and previous findings[6] question the importance of stimulus ambiguity as a determinant of projective responses, at least within the methodological limitations of these studies.

Summary

This study investigated the projective effects of fear and sexual arousal upon the rating of neutral pictures. The exposure time and the sex of the pictures were varied to ascertain the relationship between these stimulus characteristics and similar and complementary projection.

Three experimental groups of 24 male college students were formed: (1) a fear arousal group; (2) a sexual arousal group; and (3) a control group. The fear arousal conditions involved shock and threat of additional shock. The sexual arousal condi-

tions included the rating of pictures of nude women in the presence of an attractive female *E*. The *S*s of the control group rated the neutral pictures without prior experimental activity.

Every *S* of the experimental groups then rated eight male and eight female pictures which were projected onto a screen. The pictures were selected for their neutrality or ambiguity of facial expression and were presented at either a short or a long exposure time. Four scores were obtained from the *S*'s ratings of the pictures on a forced-choice adjective check list. The major findings include:

1. While there was no evidence of the similar projection of either fear or sex, the fear arousal group evidenced complementary projection of fear. These findings suggest that the occurrence and /or type of projection may be related to the particular motivation involved.

2. There was an inhibition of similar projection of sexual responses by the sexual arousal group.

3. Although the sex of the pictures significantly affected their rating, their ambiguity, as defined by the exposure time, did not.

References

1. Clark, R. A. The projective measurement of experimentally induced levels of sexual motivation. *J. exp. Psychol.*, 1952, *44*, 391-399.
2. Hornberger, R. H. The projective effects of experimentally aroused response tendencies on the rating of pictures. Unpublished doctoral dissertation, State University of Iowa, 1957.
3. Kenny, D. T., and Bijou, S. W. Ambiguity of pictures and extent of personality factors in fantasy responses. *J. consult. Psychol.*, 1953, *17*, 283-288.
4. Lindquist, E. F. *Design and analysis of experiments in psychology and education.* Boston: Houghton Mifflin, 1953.
5. Murray, H. A. The effect of fear upon estimates of the maliciousness of other personalities. *J. soc. Psychol.*, 1933, *4*, 310-329.
6. Weisskopf, Edith A. An experimental study of the effect of brightness and ambiguity on projection in the Thematic Apperception Test. *J. Psychol.*, 1950, *29*, 407-416.

THE INCOMPLETE SENTENCES BLANK: SOURCES OF VARIANCE IN RETEST RELIABILITY

MARK W. STEPHENS

Purdue University

Introduction

Churchill and Crandall [1] have reported biserial correlations of .37 (males) and .42 (females) between adjustment scores on Rotter's Incomplete Sentences Blank, administered at time of college entrance, and the criterion of later appearance *vs.* nonappearance for psychological counseling. Such a *predictive* validity, especially with an imperfect criterion, is particularly encouraging for any instrument purporting to measure "adjustment", and appears even more impressive than the *concurrent* validity coefficients of .62 and .50, for men and women, reported originally by Rotter and Rafferty [2]. Also reported by Churchill and Crandall, however, were retest reliability coefficients of the adjustment scores. For male subjects these were .43, .52, and .38 for 6-, 12-, and 36-month intervals; for females they were .54, .50, and .44. These reliability coefficients are discouragingly low; and their juxtaposition to the nearly equally high validity coefficients appears puzzling.

To account for the low reliability Churchill and Crandall suggested that much of the apparent "retest unreliability" might be due to "real" changes in adjustment

within the sample; they proposed that the first 6, 12, or 36 months of college experience might well be expected to affect students differentially in regard to their overall adjustment. Supporting this hypothesis, a considerably higher retest reliability coefficient of .70 was reported for a group of mothers, retested at a median interval of 20 months, who were participating in the research program of the Fels Research Institute for the Study of Human Development. Such an explanation, although reasonable, requires careful testing. Especially when a test has shown an encouraging degree of validity, and especially when it could serve such a valuable function as affording a valid and quantitative index of adjustment, sources of error and instability in the test and/or in the scoring thereof need to be isolated, in order to refine the test further. The present research tests the hypothesis that change in retest scores reflects "real" change in adjustment and explores other factors which may contribute to the instability of scores on retesting.

Method

The protocols of the 6-month retest groups used originally by Churchill and Crandall were made available for this study. Forty-five subjects of each sex were available, each with an original-testing protocol and a retesting protocol; data analysis was conducted separately for each sex, both because of likely sex differences in some sex-specific stems in the ISB and to afford a cross-sex replication of results. Retest reliability coefficients were computed separately for each of the forty stems. The score for a single subject on a given stem represents the degree of maladjustment reflected in his response to the stem, as determined by scoring instructions and examples in the ISB manual. The retest reliability coefficient for the stem therefore represents the degree to which each subject manifested an equivalent degree of maladjustment in his responses to the same stem on original testing and on retesting.

Seven judges then rated each stem on a 5-point scale as to the probable sensitivity of the stem in reflecting changes in adjustment during the first six months of college experience. A stem which would be expected to elicit relatively little change in adjustment was to be rated 5; a stem expected to be most sensitive to change in adjustment was rated 1. To maximize range of ratings, judges were asked to attempt to use the 5 points on the scale more or less equally frequently; no attempt was made to force a perfectly rectangular distribution of ratings, however.

The judges were all clinical psychologists or clinical psychology graduate students. In other respects they were quite heterogeneous, ranging from first-year graduate students to a Ph.D. with three years' post-doctoral experience, and from considerable training and experience with the ISB to none whatsoever.

To explore other sources of variance in the retest reliability of the separate stems, the *interrater* reliability coefficient for, and the standard deviation of scores on, each stem were computed, using the original (rather than the retest) protocols.

Results

Table 1 presents the matrix of intercorrelations of retest reliability coefficients, ratings of predicted stability, interrater reliability coefficients, and standard deviations. In all correlations of correlation coefficients, the Pearson r's were converted to z scores. The total interjudge reliability for ratings of predicted stability was .57 for males and .61 for females; correlations involving these ratings are presented first without and then, in parentheses, with correction for this attentuation.

Churchill and Crandall's hypothesis receives support, for both males and females: the correlation of .35 and .32 between retest reliabilities of stems and their *predicted* stability are both significant at the .05 level. Equally significant, however, are the correlations between *retest* reliabilities and *interrater* reliabilities, and between retest reliabilities and standard deviations. Most impressive are the correlations of .65 (males) and .73 (females) between interrater reliabilities and standard deviations. Although rater reliability of the ISB total score has never

TABLE 1. INTERCORRELATIONS OF RETEST RELIABILITIES, INTERRATER RELIABILITIES, STANDARD DEVIATIONS, AND PREDICTED STABILITY

	Rater reliabilities	SD	Predicted Stability
Retest reliabilities			
Males	.426**	.431**	.352* (.467**)
Females	.423**	.358*	.319* (.410**)
Predicted Stability			
Males	.086 (.114)	.180 (.239)	
Female	.233 (.299*)	.156 (.200)	
SD			
Males	.652**		
Females	.733**		

*$p < .05$ **$p < .01$

appeared to be a problem, being consistently in the .90s, it is impressive that what rater unreliability there is in specific stems appears to be so direct a function of variance of scores on the stems. Restricted variance of scores on a stem, then, markedly suppresses rater reliability, and either directly or indirectly suppresses retest reliability as well.

SUMMARY

This study investigated sources of instability, or retest unreliability, of the Incomplete Sentences Blank. The test has shown promising predictive validity as a measure of adjustment. However, retest reliability has been found to be relatively low, ranging from .38 to .54 for 6-month to 3-year intervals for college subjects. Since retest unreliability can be expected to limit predictive validity, identifying the sources of this unreliability can indicate means by which the validity of the test might be increased.

The primary goal of the research was to test the hypothesis that instability of test scores is in significant part a function of instability of, or changes in, adjustment itself. Computing the retest reliability coefficients separately for each of the 40 stems over a 6-month interval, this hypothesis was supported: the stems with lowest retest reliability were those which were rated by judges as being most likely to be sensitive to change in adjustment in the first 6-months of college experience. Significant correlations of .35 and .32, for men and women, were found between observed retest reliability and predicted stability of the stem.

Also significantly related to the stems' retest reliabilities were their *interrater* reliabilities and the variance of scores accruing to each stem. One expedient to boost retest reliability of the test would seem, therefore, to be to revise the scoring procedure so as to force greater variance in scores on the less reliable stems, and/or by other means to enhance interrater reliability of these stems. This procedure could, however, add spurious variance to the total score and thus reduce its validity. Such revision of scoring instructions would therefore seem best to be done with criterion data as to the validity of each stem and of each general scoring principle for the stem.

REFERENCES

1. CHURCHILL, R., and CRANDALL, V. J. The reliability and validity of the Rotter Incomplete Sentences Blank. *J. consult. Psychol.*, 1955, *19*, 345-350.
2. ROTTER, J. B. and RAFFERTY, J. *Manual for the Rotter Incomplete Sentences Blank, College Form.* New York: Psychological Corp., 1950.

DIFFERENCES IN PERFORMANCE ON THE CHROMATIC VS. ACHROMATIC H-T-P DRAWINGS*

VYTAUTAS J. BIELIAUSKAS
Xavier University

ANN R. HEFFRON
Tulane University School of Medicine

INTRODUCTION

The idea of chromatic H-T-P drawings is credited to Payne[4] who approached the problem primarily from the point of view of clinical observations. Buck suggested that "the subject's reaction to and use of color in the chromatic phase furnishes evidence of his tolerance of, control of, and reaction to additional emotion-arousing stimuli, and at the time when he is believed to be peculiarly vulnerable thereto"[2, p. 6]. The use of colored crayons instead of a pencil presumably represents a new experience to the subject and presumably influences the entire performance on the test.

This study was designed to test the influence of the color drawing medium upon the quantitative aspects of the performance by comparing the achromatic and chromatic drawings produced by the same subjects in terms of the H-T-P IQ's obtained from these drawings. The hypothesis was advanced that there would be no difference between the IQ's obtained from both sets of drawings.

PROCEDURE

Students in a metropolitan high school of an Eastern city of the U. S. A. were used for this investigation. This high school was one of the three high schools of a city with approximately 100,000 population. A sample was selected from the senior class, using the alphabetical list of names. The first 25 boys and 25 girls were placed in Group I and the next 25 boys and the next 25 girls were placed in Group II, and assumed to be representative of the class as a whole, being drawn unselectively in terms of section, academic placement or level of intelligence. Both white and negro students participated.

The subjects were administered the H-T-P using the conventional administration procedures. Group I was administered at first the achromatic H-T-P and short P-D-I, then the chromatic H-T-P and short P-D-I. For group II the order was reversed. The short P-D-I's were given in order to obtain the essential information needed for quantitative scoring. Later both groups were administered the Otis Self-Administering Test of Mental Ability (Higher Examination: Form A; 30 minutes). Both the H-T-P and Otis were scored according to regular scoring procedures described in the manuals of both techniques. The subjects were matched on Otis IQ's with results presented in Table 1.[1]

The 100 subjects were used in all statistical procedures with the exception of the analyses of variance where only 76 of these subjects could be used in matching them on three levels of Otis (high, medium, low). The subjects were eliminated where necessary by chance selection: drawing from a pile of numbers representing each subject in a given category. To test the statistical significance of variances, eight analyses of variance were computed for the IQ's from the achromatic drawings to test the influence of order of presentation, sex differences, intelligence and interaction. A t test of significance between the differences of the means of achromatic and chromatic H-T-P IQ's was computed for each group of subjects using the procedures described by Edwards[3, pp. 176-177].

---

*The authors wish to thank Dr. R. Omer Lucier for his valuable assistance in arranging of the experimental design.

[1]Tabular materials can be obtained in writing to: Dr. V. J. Bieliauskas, Department of Psychology, Xavier University, Cincinnati 7, Ohio.

Results and Discussion

The male subjects of both groups showed more significant differences between the means of the achromatic and chromatic IQ's than the female subjects.[1] The widest difference in the means of the achromatic and chromatic H-T-P IQ's for both males and females occurred on the Flaw score IQ: for males this differential was 31.52 points and for females 16.72 points, all differ significantly at .01 level of confidence. The *t* tests of significance, computed for both males and females within each of the two groups, indicated differences between the four achromatic and chromatic H-T-P IQ's significant at or beyond the .01 level of confidence.

Further analysis of our data reveals an interesting trend in these differences between chromatic and achromatic mean IQs. Both male and female subjects in both groups scored consistently higher on the first phase of the H-T-P irrespective of the type of the drawing medium used during the first phase. The analyses of variance offer interesting clues in that the *F* values show that out of seven possible sources of variation, the order of presentation is consistently significant at either .01 or .05 level. The order of presentation rather than the drawing medium was responsible for significant differences in performance since the subjects' drawings were always quantitatively better during the first phase as compared with the second phase. On the whole, female IQ's on both the achromatic and chromatic phases of the H-T-P were found to be higher than the male IQ's, although Otis IQ's were comparable for both sexes.

It was also noted that more statistical significance could be observed on the Net Weighted and Good Score H-T-P IQ's than on the Raw Score and Flaw Score IQ's. This finding seems to be in agreement with previous findings by Bieliauskas[(1)], suggesting that Net Weighted Score and Good Score IQ's on the H-T-P have higher reliability than the Raw Score and Flaw Score IQ's.

Summary and Conclusions

An unbiased and representative sample of 100 high school seniors, divided into two groups each consisting of 25 males and 25 females, were administered the chromatic and achromatic phases of the H-T-P technique in reverse order. Both groups were matched for intelligence on the Otis Self-Administering Test of Mental Ability.

Sixteen *t* tests were computed to determine the significance of differences in H-T-P IQ's obtained from achromatic and chromatic drawings. Eight analyses of variances were calculated to determine the effect of presentation, sex, intelligence, and interaction of these factors upon the quantitative scores of the H-T-P. Significant differences were found between quantitative scores of the achromatic drawings. The phase presented first resulted in higher scores without regard to drawing medium. Presentation (P) was found to be the main source of variation with the factor of sex showing some influence particularly upon the chromatic drawings.

References

1. Bieliauskas, V. J. Scorer's reliability in the quantitative scoring H-T-P technique. *J. clin. Psychol.*, 1956, *4*, 336-369.
2. Buck, J. N. *Administration and Interpretation of the H-T-P Test: proceedings of the H-T-P workshop held at Veterans Administration Hospital, Richmond, Va., March 31, April 1, 2, 1950.* West Los Angeles: Western Psychological Services, 1950. (Mimeographed).
3. Edwards, A. L. *Statistical Analysis.* New York: Rinehart, 1951.
4. Payne, J. T. Comments on the analysis of chromatic drawings. In Buck, J. N. The H-T-P technique: a qualitative and quantitative scoring manual. *J. clin. Psychol., Monogr. Suppl.*, 1948, No. 5, 119-120.

INTERRELATIONSHIPS AMONG PORTEUS MAZE TEST QUALITATIVE ERRORS

RICHARD F. DOCTER[1]

University of California, Los Angeles

Problem

A resurgence of interest in the Porteus Maze Test has resulted from publication of a procedure for scoring performance according to the expressive behavior manifested [1]. The qualitative score (Q score) is based on the assessment of penalties for violation of instructions, or for such errors as cutting corners, crossing lines, or lifting the pencil. In a number of studies [1, 2, 4, 5, 6], the Q score has been shown to discriminate between groups of socially conforming and nonconforming subjects. Almost without exception these investigators have related the Q score to the concept of impulsivity of motor performance. The present study is an attempt to examine the relationships among the Q score errors; this is believed to be a desirable first step toward conceptual validation of the Q score.

Procedure

The Porteus Maze test was given to a total of 224 subjects who were classified into the following groups: 58 delinquent adolescent boys and 60 matching nondelinquent adolescents described elsewhere [1]; 60 army enlisted men diagnosed in one of the subtypes of schizophrenia and ranging in age from 18 to 44 with a median of 24; 46 army enlisted men diagnosed in one of the categories of neuroses or character disorders and ranging in age from 17 to 43 with a median of 24.

In addition to the mazes, an intelligence test score was obtained using the following tests: for delinquents the Wechsler Intelligence Scale for Children; for nondelinquent adolescents the California Test of Mental Maturity; for all adult subjects the Wechsler-Bellevue Form I. While these measures are not entirely comparable, the means and standard deviations for the three groups do not differ significantly.

The six[2] Q score variables which were intercorrelated and factor analyzed are fully defined by Porteus [4]. Other variables included were total Q score, Maze Test quotient, and an intelligence measure. These nine scores were intercorrelated using Pearson product moment coefficients. After extraction of six centroid factors, the subsequent loadings were quite small; the centroid factors were then rotated orthogonally using Kaiser's [3] Varimax method. The correlation matrix, the centroid factors, and the rotated factor loadings are given in Tables 1, 2, and 3.

Analysis

It is apparent from Table 1 that none of the Q score errors correlate highly with each other. The error scores Wavy Lines and Lift Pencil account for the greatest share (about 70%) of the variance of the total Q score. Surprisingly, these two errors do not correlate significantly with each other. By correlating these error scores with the total Q score we are, of course, in effect correlating a score with itself plus variance added from other Q score errors. We must therefore keep in mind that the rotated factors are in part artifacts stemming from the dependence of the total Q

[1]Grateful acknowledgment is made for computational assistance by Mr. Ray Perkins. I also wish to express gratitude for use of the SWAC electronic computer operated by Numerical Analysis Research at the University of California, Los Angeles, and supported by the Office of Naval Research. The computer programs used for correlation, factor extractions, and Varimax rotations were developed by Dr. Andrew L. Comrey. Some of these data were collected while the author was on active duty in the United States Army.

[2]The Q score error Last Third was excluded from the factor analysis because fewer than five percent of the subjects committed this error and the resulting distribution of errors was extremely limited. The Q score errors Year V and Year VI were excluded because these errors can not be committed by adults as they must begin the test with the Year VII maze.

TABLE 1. INTERCORRELATIONS AMONG SIX PORTEUS MAZE QUALITATIVE SCORE ERRORS, TOTAL Q SCORE, MAZE TEST AGE, AND INTELLIGENCE

Variable		1	2	3	4	5	6	7	8	9
First Third	1		01	07	00	12	13	20	−34	−07
Cut Corner	2			29	−07	07	08	09	−05	−04
Cross Line	3				−05	19	02	16	−23	−04
Lift Pencil	4					05	10	63	04	09
Wavy Lines	5						25	60	−15	−17
Wrong Direction	6							27	05	−10
Total Q Score	7								−16	−04
Maze Test Age	8									11
Intelligence	9									

score upon the sum of the Q score errors. Factors I and III are simply another way of expressing the major variance of the Q score attributable to Lift Pencil and Wavy Lines. Factor II can be attributed to the Maze Test Age and should also be considered artifactual. This factor has high loadings on Maze Test Age and First Third

TABLE 2. CENTROID FACTORS

	I	II	III	IV	V	VI	h^2
First	36	−24	−32	−28	20	−21	45
Cut	26	−27	30	26	−14	−07	32
Cross	36	−33	16	23	07	19	36
Lift	42	46	−39	26	−26	09	68
Wavy	58	11	30	−33	10	24	62
Wrong	34	06	12	−17	−09	−21	21
Q Score	83	47	−09	−07	−08	23	98
T A	−26	39	23	32	13	−13	41
I Q	07	13	−07	30	28	−09	20

but part of the definition of a First Third error involves the act of a maze failure which reduces total Test Age score. Factor IV is defined by Cut Corners and Cross Lines, both of which involve compliance with instructions and careful motor control. Whatever kind of motor performance may be involved in the commitment of these errors, it is certainly of a different kind than is involved in the drawing of Wavy lines. Factors V and VI are of no importance relative to our concern with relationships among Q score errors.

TABLE 3. VARIMAX ROTATIONS

	I	II	III	IV	V	VI	h^2
First	01	−64	06	02	09	−16	45
Cut	01	01	01	55	00	−12	32
Cross	00	−15	16	55	07	09	36
Lift	81	02	04	−06	11	−07	68
Wavy	08	−10	73	12	−09	−22	62
Wrong	09	−09	17	05	−05	−40	21
Q Score	70	−18	63	12	09	−19	98
T A	−03	53	09	14	32	−05	41
I Q	08	02	−03	05	43	04	20

SUMMARY

The Porteus Maze test and an intelligence test were administered to 224 subjects; the mazes were scored both for Test Age and for qualitative score (Q score). Error scores contributing toward the Q score plus certain other variables were intercorrelated and factor analyzed. The variance of the Q score can not be accounted for by a single factor; Lift Pencil and Wavy Lines contributed very heavily to the total Q score but did not correlate with each other. The error First Third related more

closely to the Maze Test Age than to the total Q score. Cut Corner and Cross Lines defined a separate factor which was independent of both total Q score and Test Age. It would appear indefensible to interpret the Q score as indicative of a single trait such as impulsivity or impulse control.

REFERENCES

1. DOCTER, R. F. and WINDER, C. L. Delinquent vs. nondelinquent performance on the Porteus Qualitative Maze Test. *J. consult. Psychol.*, 1954, *18*, 71-73.
2. GRAJALES, M. C. Porteus' Qualitative Maze Test as a measure of delinquency. Unpublished master's thesis, Fordham University, 1945.
3. KAISER, H. F. The Varimax criterion for analytic rotation in factor analyses. *Psychometrika*, 1958, *23*, 187-200.
4. PORTEUS, S. D. *The maze test: recent advances.* Palo Alto, Calif.; Pacific Books, 1955.
5. PURCELL, K. A note on Porteus maze and Wechsler-Bellevue scores as related to anti-social behavior. *J. consult. Psychol.*, 1956, *20*, 361-364.
6. WRIGHT, CLARE. The qualitative performance of delinquent boys on the Porteus Maze Test. *J. consult. Psychol.*, 1944, *8*, 24-26.

AVA AND THE KESSLER PD SCALE AS MEASURES OF PASSIVE-DEPENDENCY

PETER F. MERENDA AND WALTER V. CLARK AND SYDNEY KESSLER

Walter V. Clarke Associates, Inc.
E. Providence, R. I.

Los Angeles, Cal.

PROBLEM

The Kessler PD Scale (KPDS) [2] is a personality inventory measuring a passive-dependent trait cluster. The Activity Vector Analysis (AVA) [1] is a self-concept personality assessment instrument yielding profiles which may be interpreted as indicative of passive-dependent tendencies. This study was conducted to determine the degree to which the KPDS and AVA are both measures of passive-dependency.

METHOD

Subjects. Ss were 181 adult males and 99 adult females. The great majority were undergraduate students attending a large municipal college, the remainder consisting of nonstudent adults residing in the area surrounding the campus. Age range for females was 17 to 62, with median age of 25. Age range for males was 17 to 57 with median age of 22.

Procedure. The AVA and KPDS were administered simultaneously to the Ss. Scoring on the KPDS is accomplished by counting the number of passive-dependent responses made to 112 statements appearing on the scale. The author indicates the raw score of 40 as the critical point on the passive-dependency continuum. Profile scores on the AVA are obtained by counting the number of words checked in each of four vectors. These raw scores are converted to ordinary standard scores which are transformed in the individual vectors to deviation scores relative to the profile mean.

Statistical Analysis. Separate analyses were made for male and female subjects since there might be differences in the relative incidence of passive-dependency among males and females. Accordingly, the data of the study were treated so as not to be contaminated by any sex differences with regard to the trait under investigation should such differences exist.

A modified discriminant analysis technique was used following Wert [3]. The male and female samples were each dichotomized at the critical point of raw score

of 40 on the KPDS. Those Ss whose KPDS scores were 40 or over were classified into the passive-dependent category and scores under 40 were classified in the nonpassive-dependent category.

The set of discriminant variates was the set of deviation scores on the 4-factor AVA profiles. The analysis of maximum separation yields optimal weights for each variate. The multiple biserial correlation and discriminant function between the personality profile and the dichotomized criterion were calculated as were the tests of significance for these statistics. The data also lend themselves to easy calculations of intercorrelations among the predictor variables and between each of the predictors and the criterion.

Results and Discussion

Table 1 presents the distribution statistics for AVA and KPDS scores. The data of the table are organized to show the means and standard deviations of the AVA profile scores and scores on the KPDS for subjects classified as Passive-Dependent (PD $\geq$40) and nonpassive-dependent (PD $<$ 40). The same statistics are given for the total sample.

Male Sample. Inspection of these data reveals that a relatively small portion of the sample was identified as passive-dependent according to the criterion of the study. Only 31 of 181 males possessed scores on the KPDS of 40 or above. The mean KPDS score of these 31 Ss was 49.48; the mean KPDS score of the remaining 150 Ss was 23.05.

The personality profile data reveal that as a group, the Ss show predominant tendencies toward self-initiating, socially confident, energetic, and independent behavior. This inference is drawn from the average personality profile for the total sample in Table 1 which shows relatively high mean scores on AVA factors 1 and 2 (+5.53 and +2.17 respectively) and relatively low mean scores on AVA factors 3 and 4 (−3.97 and −3.58 respectively). These statistics serve to confirm the findings with the KPDS that the sample is composed of persons who are not disposed toward passive-dependency.

The tendency of the AVA and the KPDS to commonly identify the passive-dependent individuals in the sample of 181 male subjects is suggested by the other data of Table 1, broken down by performance on the KPDS. It is noted that those Ss who scored above the critical point of 40 on the KPDS possessed as a group personality profiles which showed higher scores on the passive (V-3) and dependency (V-4) vectors of the AVA, as compared to those with lower KPDS scores. These differences are in the hypothesized direction. These tendencies are confirmed by the data of Table 2 which show that a substantial and highly significant relationship was found to exist between the AVA profiles and KPDS scores.

The negative correlations obtained between the AVA vectors of aggressiveness and sociability show that these factors individually are negatively related to measures of passive-dependency as measured by the KPDS. The positive correlations between emotional control and social adaptability show these factors to be positively related to passive-dependency. The multiple correlation for the AVA profile set was found to be .36. The multiple biserial correlation obtained by discriminant analysis techniques for the passive-dependency dichotomy based on the KPDS cut-off of 40 was found to be .50. Both values are statistically significant beyond the .001 level of significance. These findings reveal that the AVA and the KPDS overlap significantly in their measurement of passive-dependency in males.

Female Sample. Table 1 presents similar distribution statistics for the female subjects. As suspected, sex differences appear to exist with respect to passive-dependency. Twenty-seven per cent of the females attained scores on the KPDS which were classified as indications of passive-dependent temperament, whereas only 17% of the males achieved scores as high as 40 on this scale. Sex differences are also suggested by the examination of the AVA personality profile data of Table 2.

TABLE 1. DISTRIBUTION STATISTICS OF AVA AND KPDS SCORES FOR PASSIVE-DEPENDENT AND NONPASSIVE-DEPENDENT SUBJECTS

Factors	Passive-Dependent (KPDS ≥ 40)				Nonpassive-Dependent (KPDS < 40)				Total Sample			
	Males (N = 31)		Females (N = 27)		Males (N = 150)		Females (N = 72)		Males (N = 181)		Females (N = 99)	
	$\bar{X}$	σ	$\bar{X}$	σ	$\bar{X}$	σ	$\bar{X}$	σ	$\bar{X}$	σ	$\bar{X}$	σ
Aggressiveness	+3.53	9.21	+0.21	8.92	+5.94	9.90	+3.86	7.41	+5.53	9.83	+2.86	8.02
Sociability	−0.44	6.88	−1.67	6.48	+2.71	7.35	+4.81	7.47	+2.17	7.37	+3.04	7.77
Emotional Control	−1.54	8.14	−2.84	6.48	−4.47	8.05	−4.34	6.97	−3.97	8.14	−3.81	6.87
Social Adaptability	+0.27	6.24	+2.19	7.34	−4.37	6.08	−4.25	5.87	−3.58	6.36	−2.50	6.92
Passive-Dependency	49.48	10.66	50.33	9.25	23.05	8.27	25.32	8.20	27.62	13.24	32.14	14.01

TABLE 2. RELATIONSHIPS BETWEEN AVA AND KPDS SCORES IN MEASURING PASSIVE-DEPENDENCY

Variates	Correlations[1]					(PD-NPD) Mean Diff		Validity Statistics	
	V-1	V-2	V-3	V-4	KPDS	Males	Females	Males	Females
V-1 (Aggressiveness)		+.11	−.64	−.57	−.32	−2.41	−3.64	R = .36*	R = .66*
V-2 (Sociability)	−.03		−.54	−.63	−.42	−3.15	−6.49		
V-3 (Emotional Control)	−.67	−.48		+.40	+.16	+2.94	+1.50	R_{bis} = .50*	R_{bis} = .84*
V-4 (Social Adaptability)	−.57	−.51	+.43		+.48	+.464	+6.44		
Criterion (KPDS)	−.14	−.26	+.18	+.34					

[1]Correlations above diagonal are those for females (N = 99); those below diagonal are for males (N = 181)
*Significant beyond .001 level

The females classified as passive-dependent possessed, on the average, AVA profiles with substantially lower scores on the sociability factor and higher scores on the social adaptability factor than did the males. This finding suggests that dependency in women is accompanied by socially withdrawn behavior to a greater extent than appears in males. It is the integration of these two factors which contribute most heavily to the exceptionally high multiple R of .66 obtained for the personality set against the criterion of KPDS scores and multiple biserial R of .84 against the dichotomized criterion of passive-dependency vs. nonpassive-dependency.

Summary

The Kessler PD Scale and the Activity Vector Analysis were administered simultaneously to a group of 181 adult males and 99 adult females as measures of a passive-dependency trait cluster. Our data show that passive-dependency as defined according to the scoring key of the KPDS is also measured to reasonably high extent by the AVA. Differences between males and females become apparent with the later showing greater tendency toward passive-dependent behavior both with regard to frequency and to extent of passive-dependency. A substantially greater proportion of the female sample was classified as being passive-dependent by performance on the KPDS. On the AVA, sharper discriminations between the high and low groups on the passive dependency continuum were evidenced. Finally, it is apparent from the results obtained from both samples that passive-dependency as measured by the KPDS is a reflection of the inability of the person to be socially-outgoing, but at the same time wanting to conform with and adapt to the accepted rules of society.

References

1. Clarke, W. V. Construction of an industrial selection personality test. *J. Psychol.*, 1956, *41*, 379-394.
2. Kessler, S. The Kessler PD Scale. Los Angeles, 1953.
3. Wert, J. E., Neidt, C. O. and Ahmann, S. *Statistical methods in educational and psychological research.* New York. App.leton-Century-Crofts, 1954, 256-281.

THE EFFECTS OF TWO SIMULTANEOUS COGNITIVE AND AFFECTIVE STIMULI ON A GROUP OF CHRONIC SCHIZOPHRENIC PATIENTS WITH SUICIDAL IDEATION*

EMANUEL STARER

Veterans Administration Hospital, Northport, New York

Introduction

This investigation explores the effects of affective stimulation on psychomotor performance in a group of chronic schizophrenic patients with suicidal ideation or overt suicidal activity. Malmo, Shagass, Bellanger and Smith[1] in a study involving 76 psychiatric patients and 21 normals as a control reported that psychoneurotic disorders involve defective motor regulation, manifested by abnormally increased motor disturbance under any stressful condition. Starer[2], using the method of two simultaneous cognitive and affective stimuli on a control group of 15 student nurses compared with an experimental group of 15 female psychiatric patients, found that differences in the performance between the two groups included psychomotor disorganization, delayed reaction time, and behavioral disorganization in the psychiatric group when compared with the control group.

*The statements and conclusions published by the author are the results of his own study and do not necessarily reflect the opinion or policy of the Veterans Administration.

Method

The control group consisted of 24 chronic schizophrenic patients, mean age 36.8 years, S.D. 7.8, average length of hospitalization 6.8 years, and an average of 10.2 years schooling. None had any previous history of suicidal ideation or overt suicidal activities. The chief administrative problem with the control group lay in their tendency to try to leave the hospital without permission.

The experimental group consisted of 25 chronic schizophrenic patients, mean age of 36.9, S.D. 10, average length of hospitalization 5.4 years, with an average of 11 years of schooling. Each had a history of previous overt suicidal attempt on at least one or more occasions or had expressed overt suicidal ideation. No learning difficulty or uncorrected visual difficulty was present in any of the subjects.

The subjects in both groups were given a series of 15 designs to copy on 8½ x 11 white paper. These designs consisted of a circle with the number of each design in the center. On top of this circle was a rectangular-like figure attached to the circle with a series of dots inside arranged in a triangular shape. A wavy line was attached to the end of the right hand side of the rectangular-like figure and extended to the bottom of the circle. This basic figure was rotated at different angles to produce seven other figures, exactly the same except for the angulation. The subjects were all provided with two well-sharpened pencils with erasers. The standard instructions were: "Please copy the following designs as well as you can on the paper." If the subjects wanted more paper, they were given it quickly.

Within a period of five days the same designs were given to the subjects in the identical order. The instructions this time were: "Please listen to the following instructions. You are going to be given a series of designs to copy as well as you can on the paper. While you are copying them, you will hear me speaking to you. Continue copying the designs no matter what you hear. Do not answer any questions that will be asked. Do not stop copying the designs at any time. Are you ready?" (Pause of 10 seconds.) "All right, now begin copying the designs." (Pause of 5 seconds.)

Two sets of questions were asked of each group of subjects during the copying of the designs. On designs 3, 5, 7, 8, 9, 13, and 14, questions involving suicidal activity were asked. On the remaining designs questions of a relatively non-affective nature were asked. Examples of such questions included were as follows: Do you enjoy watching television? Do you like baseball games? Do you enjoy playing checkers or chess? Do you like to spend time reading historical novels? Do you like to swim? Do you enjoy football games?

Examples of questions involving suicidal ideation or activity included: Do you want to harm or kill yourself? Do you feel like doing away with yourself? If you found a gun, would you shoot yourself? If you found a knife or razor blade, would you cut yourself? Do you feel like jumping off a bridge? Do you want to die?

The copied designs in both groups were compared before and after affective stimulation in order to determine whether they were the same, improved, or worse than the original performance.

Results and Discussion

All 24 subjects of the control group showed improved performance after affective stimulation as compared with 11 subjects in the experimental group improving while 14 were worse. The Chi Squares showed these differences between the groups are significant at the one per cent level. The psychomotor disorganization in the experimental group was at times very severe. Rotations, distortions, perseverations, inappropriate planning, and refusing to complete the test were present. At other times, marked affective display in the form of inappropriate laughing or smiling was noted in some of the patients in the experimental group. This was in sharp contrast to the control group where this behavior was absent.

The important question is raised why 11 of the experimental group performed approximately as well as the control group. This may be partly explained by the fact that suicidal ideation was no longer present in some of the group on special observa-

tion. In fact, five of the 11 patients in this group were transferred to other wards after being removed from special observation group status. Several of the remaining expressed the wish to die but indicated that they would not harm themselves by their own hand.

One unusual situation developed where one of the patients openly expressed suicidal ideation but no difference before and after affective stimulation was noted in the design copying. It was observed that he had his left hand tightly clenched in a fist-like position throughout the test. It is conceivable that the psychomotor test tension which would ordinarily be manifested in the task performance was thus artificially prevented from appearing.

Summary

This investigation is an exploratory attempt to study the effects of affective stimulation on psychomotor performance employing a group of chronic schizophrenic patients with suicidal ideation or overt suicidal activity. A control group consisted of 24 chronic schizophrenic patients, mean age of 36.8 years with a sigma of 6.8 years. No previous history of suicidal ideation or tendency appeared in this group.

The experimental group consisted of 25 chronic schizophrenic patients, mean age 36.9 years with a sigma of 10, and average length of hospitalization 5.4 years. Each one of these patients had a history of suicidal ideation or tendency on at least one occasion. Differences between the two groups included psychomotor disorganization and behavioral disorganization in the experimental group when compared with the control. The technique of simultaneous impingement of stimuli appears to merit careful consideration and further exploration in the detection of patients with suicidal ideation or suicidal tendencies.

References

1. Malmo, R. B., Shagass, C., Belanger, D. J. and Smith, A. A. Motor control in psychiatric patients under experimental stress. *J. abnorm. soc. Psychol.*, 1951, *46*, 539-547.
2. Starer, E. The effects of two simultaneous cognitive and affective stimuli. *J. clin. Psychol.*, 1952, *8*, 402-405.

OPERATIONAL PSYCHOLOGICAL REPORT WRITING

FREDERICK C. THORNE

Brandon, Vermont

The Goals of Scientific Reporting

Modern science faces the problem of maintaining open channels of communication in spite of the increasing complexity of subject matters. Perhaps the most important problem is the proliferation of new vocabularies in every subdivision of science to the point where only a semantic specialist can differentiate all the meanings involved. Psychological science, particularly, is greatly handicapped by an undue proliferation of vocabularies related to systems and schools to the extent that even well trained basic scientists find it difficult to comprehend the content of specialty journals. *Obfuscation* is the cardinal sin of scientific reporting. Our goal must be to make the clearest most comprehensible presentation in simplest terms, using the smallest possible number of words, dealing only with relevant issues and avoiding undue generalization and speculation. The aim is to clarify and not to make obscure.

Although the main purpose of scientific journalism is to preserve and communicate knowledge, we must not overlook the secondary issues (such as the psychological characteristics of the average journal reader) which actually may determine how well specific objectives are achieved. Our editorial impression is that it typically re-

quires 10-20 years for a new concept presented in the literature to achieve general assimilation and then only if a large number of secondary considerations happen to have been satisfied. The current literature has assumed such vast proportions as to recommend seriously the execution of systematic scientific studies to determine how knowledge can be condensed and communicated in capsule form such that the average scientific intellect can grasp it. Your editor personally must confess that he cannot understand much of what appears in contemporary psychological journals without an inordinate amount of study clearly unjustified by the importance of the material presented. With this in mind, we have worked out an operational format for this Journal which we believe can be readily comprehended by any psychologist with rudimentary basic science training.

The Operational Outline

Ordinarily, the topical outline of the paper should be subsumed under the headings: Problem, Method, Results, Discussion, Summary and Bibliography. It is our policy to discourage authors from adding extra sections unless there is clear indication for discussing special issues in more detail. Starting with the title of the paper, everything should be stated in concise operational terms. Every unnecessary paragraph, sentence, phrase, words or even letters should be struck out. All key terms should be defined operationally. The goal is to keep the level of operational reporting as close to the actual experiential level as possible.

Problem. The author should give a terse operational statement of the problem in the first sentence. Ideally, this introductory statement should indicate not only the hypothesis under study but also some indication of the methods and subjects used. This can be achieved only by an operational description of what the study is actually about.

Ordinarily we do not encourage authors to include an extensive historical summary under the statement of the problem. This is a hangover from the advanced experimental psychology course in which students are required to include an extensive review of the literature in their reports of experiments in progress. In scientific journalism it is assumed that the professional reader is familiar generally with the field and does not require repetitious summarization with every new paper. Authors should make reference only to studies which are directly pertinent to the specific issue under study, and should not include references which have only tangential or general significance. Usually, the background of the problem may be stated in a few terse sentences summarizing pertinent issues. Where the issues are controversial, it is sufficient to make a simple statement that certain studies support one conclusion while others support another conclusion. Where available, the reader may be referred to general surveys or reviews of the field.

Usually, the statement of the problem will conclude with a statement of the hypotheses under study. In general, hypotheses should be stated in *apriore* rather than in *a posteriore* terms. The purpose of a study is to discover *what will result when certain things are done,* and the statement of the problem and hypotheses should reflect this open ended question. Some authors appear to be arguing after the fact when they postulate a large number of hypotheses tailor-made to fit data which have obviously already been obtained.

Many authors appear to seek justification, support and prestige for their study by attempting to connect it up with the work of established authorities in the field. Editors recognize this subterfuge immediately by noting generalized references in the bibliography to books by noted authors. Polite as it may be for authors to tip their hats to Freud, etc., in passing, we rigorously strike out such references unless they bear specifically on the problem at issue.

Method. Ordinarily the statement of the method will involve subheadings referring to Subjects, Materials, Procedures and Methods of Analysis. The writer should follow this order so that the reader may understand the naturally unfolding operational sequence underlying the experimental design.

Unless some new materials or methods are being reported, it is unnecessary to describe in detail the nature and mode of administration of tests and methods which are well known to the profession. Similarly, it is not necessary to describe the rationale of statistical methods in detail nor to include detailed computations. It is assumed that professional readers are familiar with standard statistical methods and that authors have performed computations correctly.

Where new tests and methods are being reported, the author should describe them in sufficient detail so that readers can understand their nature and use them. It is useless even to mention a new test unless the author intends to make it available to readers.

Results. Multiple reporting of the same data in text, tables and graphs should be avoided. Instead, the author should choose the medium which presents the data most clearly. Where tabular presentation is crystal clear, it is undesirable to repeat the same numbers in the text.

Generally, where there are large amounts of data, it is necessary only to cite significant findings and omitting nonsignificant figures. Secondary tabular materials and data may be referred to the American Documentation Institute for the benefit of the few readers who may wish to analyse the basic data for themselves. The same data should not be reproduced in several places. Thus, in tables of correlation data, the same figures need not be reproduced in upper and lower halves of the table.

Graphic presentation of data is usually a luxury, requiring both expensive cuts and uneconomic use of space. Although graphs may illustrate general trends more clearly, the professional psychologist usually prefers statistical presentations giving exact figures or values.

Data should be presented in the smallest number of tables or graphs possible. When successive tables have the same headings on X or Y axes, the author will usually find that they can be combined. Do not include material in tabular form which could as well be presented in one or two sentences in the text. This refers particularly to one or two line tables. Tables should be constructed according to the proportions of the journal page with the longer X axis presented horizontally and the shorter Y axis vertically wherever possible.

Discussion. The Discussion should be used to elaborate points of positive significance rather than to speculate on the significance of negative findings and what might have been. Too many authors use the discussion to simply reformulate and repeat significant points presented under Results. Restatement of significant findings should be limited to the Summary.

Many authors seem to feel a compulsion to rationalize negative findings particularly when they confute some authority. They seem to feel that it is impolite to present evidence contradicting some authority, so they attempt to rationalize the differences as being caused by slight differences in samples, procedures, etc. Let negative results speak for themselves without speculation as to causes.

Most authors also seem to suffer from the compulsion to point out that the study should be repeated with larger samples, more rigid controls, etc. Such obvious generalities are so apparent that they do not require routine repetition in the concluding paragraph of the Discussion in every paper.

Similarly, the discussion should not be used as a place for staking out priority claims of having first thought about what course future developments may take. Authors should stick to their own results and refrain from giving gratuitous advice concerning how other workers should follow up the issues.

Summary. Whether the Summary should consist of a complete abstract summarizing problem, method, results, discussion and conclusions, or should consist primarily in a statement of findings and conclusions, should be determined by the length of the paper and the magnitude of the problem. With long papers where complex problems, methods and results are involved, it is desirable for the summary to be a complete abstract of all these topics. With shorter papers, less than one page in length,

where the reader can take in all topics at a glance, the Summary may properly include only the conclusions.

In any case, the Summary should provide the reader with a complete and faithful statement of what the problem and results are about in operational terms.

Operational Definitions

All basic concepts and variables should be defined at the beginning of the paper under the statement of Problem in strictly operational terms. If such semantically abstract terms as *psychological rigidity* are used, they should be operationally defined in terms of the operations whereby they are presumably being dealt with.

Much confusion is introduced by the effort of writers to relate their findings which have specific operational referents to more general theoretical constructs and psychological systems. The concept of "intervening variables" is most commonly used to bridge the gap between specific experimental findings and more general theoretical constructs. What is the operational definition of "intervening variable"? "Intervening variables" in truth are simply "unknown factors" and involve a semantic subterfuge if anything more is implied. At the current stage of basic science evolution, it is unprofitable to speculate about the "unknown" or "uncontrolled" factors which exist in every situation. One may mention "intervening variables" simply to indicate awareness of unknown factors which must be presumed to exist. Or they can be omitted under the presumption that everyone accepts their existence.

Operational definitions largely render unnecessary higher level semantic abstractions when attempts are made to define concepts verbally. Some authors have a penchant for abstract titles such as "Personality Structure in Terms of the Concept of Psychological Rigidity" and then require several pages of semantic juggling and verbal dialectics to define what is meant by *personality structure* and *psychological rigidity*. All this may be avoided by labelling the paper "A Correlation Study of MMPI and Water Jar Test Results in College Sophomores". Operational definitions render unnecessary high leval semantic abstractions. They encourage the empiric experiencing of the same methods and data by the reader or any who wish to repeat the experiment.

Vocabulary

Basic English should be the basic tool of the scientific writer. Literary sophistication is revealed by the ability to communicate meanings in the simplest most clear terms. Conversely, pedantry obscures scientific reporting by introducing excessive displays of overcomplicated terminology satisfying only the author's Ego needs. Many observers have commented on the tendency of scientists to complicate reporting by use of esoteric polysyllabic terminology to give the illusion of sophistication and higher knowledge. Historically, medicine has been particularly guilty of using obfuscating vocabularies to create a secret proprietary language to impress the layman and give the impression of greater knowledge than actually exists. Among the devices used to create the illusion of scientific sophistication are the use of foreign language words, the coining of esoteric terms, and the invention of alternative scales or measures.

Citing References

References should be cited *only* when they refer to the specific problem at issue. Often it is sufficient in one sentence to simply state that certain studies support the validity of a finding while others disagree. Let the reader make his own search of the literature if he so desires. Let everyone evaluate the evidence for himself by reference to original sources.

This Journal continues to follow the old style of citing references by number only, and *not* following the new APA style of citing the date of the reference. This new style is awkward and clumsy particularly when a large number of references are cited in one place, e.g., Smith (1956); Mumpower, Dingle & Bonk (1909); Wax-

wing, Pickerell & Gashouse (1918); Fink et al (1962 a,b,c) . . . . . Such a style of presentation is not esthetic, efficient or convenient for the reader.

Titles

The goal of terse operational reporting should be extended to the wording of titles, heads and subheads in text, tabular and graphic materials. The main title of the paper should be operationally descriptive exactly of what the paper is about. Titles should not be phrased in generalities which bear little relation to the specific problem.

Topical headings should be arranged in logical outline form but usually should not be numbered. Except for lower case reference citations in parentheses, numbers should be used in the text only to refer to data. The only exception to this is a clearly numbered outline.

Special Rulings or Markings

Every editor has his own style for designing the layout or format of the article and prefers to use his own markings to indicate to the printer what is desired. In order not to interfere with the editorial marking system, authors should not use any special underlinings, rulings, undue capitalizations, etc., other than the standard proof-reader's symbols placed in the margin. Underlining should be used only where the author desires italics. The author should not indicate rulings in tables, leaving this to the editor. Do not write anything in all capitals—just use capitals in normal positions and everything else in lower case.

Authors should examine a recent copy of the Journal to which the article is to be submitted and get some idea of the format and style. This will provide some indication for how headings, titles, indentations, etc., should be arranged to conform to the established style of the journal This is particularly important in the citing of references in the Bibliography. A surprisingly large number of authors do not cite references properly.

Scientific Literary Style

Scientific reporting requires a specialized style involving the qualities of formal, streamlined exactness. In its final form ready for publication it should represent a terse abstract, comprehensive and yet stripped of all nonessentials. How condensed the abstract should be will be a function of the importance of the problem and the positive significance of the results.

Many authors seem to feel the compulsion to act out several roles simultaneously, showing by their literary style that they are humorous, artistic, sophisticated, humanistic, democratic, etc. Thus we encounter the flossy *bon mot*, the well-turned phrase, the elegant redundancy, the compulsive repetitiousness and the pedantic attention to nonessentials which serve almost as literary trademarks for some authors.

Anatole France is reported to have originated the following outline for correcting a manuscript:*

1. Weed out the who's, which's and whom's.
2. Shorten sentences. Strike out phrases linking sentences. Make sound transitions of thought.
3. Watch the order in sentences; recurrence of the same word (s) means that the sentence should be rewritten, not that a synonym should be used.
4. Take out useless adjectives.
5. Clip out the pastry, the adventitious and the redundant.

Fig. 1 presents some examples of unacceptable style which we have culled from submitted manuscripts.

*Reproduced from editorial comment, *Modern Medicine*, 1960, 28, p. 23.

FIG. 1. EXAMPLES OF UNACCEPTABLE AND ACCEPTABLE STYLES OF REPORTING.

Unacceptable	Acceptable
The purpose of this study is to investigate	This study investigated . . . .
Inspection of Table 1 reveals	Table 1 reveals . . . .
The first phase of this study will be concerned with testing whether or not	The first phase of this study tests whether . . . . .
The subjects of this study consisted of	Our Ss were . . . . .
It is noted from the data that whites and	Whites and . . . . .
The results reported by Wechsler are in agreement with those obtained by Smith	Wechsler's results agree with Smith's . . . . .
It was found that the left hemiplegic patients . . .	The left hemiplegics . . . .
There are four WAIS subtests which require	Four WIAS subtests require . . . .
. . . . is more difficult than the other subtests for this group.	. . . . is more difficult than the other subtests as noted by Garfield.
This result has already been noted by Garfield . . .	This study examines the importance of having . . .
This study is an examination of the importance of having . . .	Please note . . . . .
We would like to call attention to the fact . . .	These results do not agree with Dim and Dimmers' finding that the means of groups A and B differed.
In other words, our results tend to contradict the findings of Dim and Dimmer who found that the mean of group A differed from the mean of group B.	The means of Groups A and B were 15.09 (SD 3.7) and 13.52 (SD 3.4).
The mean of group A was 15.09 and the standard deviation was 3.7. The mean of group B was 13.52 and the standard deviation was 3.4.	(Even existential concepts can be communicated more fluently and validly using operational terms.)
Experiencing sequences and succession of whatnesses is experiencing hownesses . . . . It is only by experiencing our whatnesses and hownesses in their conflicting and cooperating aspects do we experience more of our wholeness . . .	

The use of footnotes is to discouraged wherever possible. Footnotes are not only expensive to insert but also cumbersome to handle and distracting to the reader.

LENGTH OF THE PAPER

Many authors seem to feel that the length of the paper should bear some relation to the amount of work invested in it and the importance of the problem. Long papers are assumed to bear more prestige because only important contributions would be assigned so much space. Scientific reporting no longer can afford to indulge such egoistic illusions. Actually, the most important findings of the most complex research often are reportable in simple formulas or in less than one page presentations. How much better for science in general, and readers in particular, if such research is presented in compendium form instead of in a long book.

Our editorial impression is that the weaker the paper, the more space the author seems compelled to use to rationalize the lack of significance of his findings. Our policy is to require authors reporting essentially negative results to condense their reports to essentially outlined abstracts.

Too often the manuscript resembles a student's report for an advanced experimental course where it is required that pertinent literature be reported in detail and all possible alternatives discussed extensively. While such length may be desirable for the professional development of the student, it has no place in scientific reporting. It is unnecessary to cite the literature in great detail. Just give the essential findings or disagreements in a few terse sentences. It may be assumed that professional readers are familiar with the literature and do not require intensive briefing with every paper. It is enough to cite the sources where such information can be obtained.

In general, papers presenting positive results can command more space than those reporting essentially negative results. Many authors appear to feel guilty on reporting negative results, particularly when they conflict with some authority, and so they engage in lengthy speculations as to how the differences might be resolved. Let the data speak for themselves and the cards fall where they may.

Serial publications of the same research are discouraged because they unnecessarily lengthen the presentation through use of multiple and repetitious citing of the same facts or data. One comprehensive paper is much more efficient than several partial presentations.

Publication Priorities

Every editor must recognize publication priorities determined by the general significance, scientific return and economic investment in any manuscript with relation to order of acceptance and publication, position assigned, space assigned and equitable assignment of publication costs to an author. At stake are the factors of (a) the editor's evaluation of the importance of the ms., (b) reader preferences concerning the nature and style of materials they wish to see reported, and (c) the author's preferences concerning length and style of presentation. It may be assumed that every author desires the most magnificent publication style available. And it must not be forgotten that if the subscribers don't like what is in a Journal they won't buy it. Ultimately, the editor must assume complete responsibility concerning the decisions of what and how to publish.

In processing any manuscript, this editor early reaches a decision concerning how much space and priority any manuscript deserves to command. We then evaluate the ms. to determine how much of its features and length represent excess space utilized by the author's preference over and above its assigned priority. Our policy is to charge authors with excess space if they do not desire to condense the ms. to suggested limits. Some ms. do not command high enough priority to be published except by author's subsidy.

It would be very helpful if authors would achieve a realistic appraisal of the priority of their own manuscripts and tailor the style of presentation accordingly. Thus, it is unrealistic to submit a masters thesis 10-12 pages in length when only the exceptional masters thesis commands priority for even minor note publication. Similarly, detailed publication of doctoral dissertations are not encouraged unless strikingly significant new results have been obtained. Such theses and dissertations usually are to be regarded as investments in the author's own personal development and have limited general significance. Promotional reports, more or less blatantly advertising a new test (usually marked by a reference to where it may be obtained, cost, etc.) do not have high priority.

Publication Costs

If realistic cost accounting was performed to discover not only manufacturing, production and distribution costs, but also the true value of editorial services usually rendered gratis to authors, it would be found that the actual per page cost ranges from $30.00 to $50.00 for any of the larger journals. Depending upon the length of the manuscript and the number of editorial consultants utilized, authors receive expensive consultations relating to the validity of the research design, the significance of the results, and the best mode of presentation. If such services were compensated at realistic levels, they may amount to several hundred dollars for editorial consultation charges alone if such were charged at $20-$25 per hour which they are worth. This is particularly true in cases where a manuscript is read by several consultants, undergoes several revisions and much correspondence. A manuscript which has been around the circuit of journals, being rejected many times, may actually represent a wasted investment of many hundreds of dollars and hours of valuable time by consultants. Authors must use their consciences concerning how much priority they can realistically demand, particularly with oft-rejected battered manuscripts.

Editors are confronted with difficult decisions concerning how much priority and space to assign to the multitude of manuscripts competing for acceptance. Acceptance of a manuscript literally means a gift of several hundred dollars to the author of a long paper. Whether publication costs are ultimately subsidized by a professional organization or by subscribers, the author is getting something paid for by other people's money. Authors have an obligation to conserve resources either by condensing articles submitted, or by assuming part of publication costs for excess space utilized. Research budgets should include publication subsidies which usually are only a small fraction of research costs. Authors desiring special services such as the publication of low priority manuscripts or utilization of excess space beyond minimum reporting requirements should plan to assume such added costs.

THE UTILITY OF THE PROGRESSIVE MATRICES (1956 EDITION) WITH NORMAL AND RETARDED CHILDREN[1]

LESLIE F. MALPASS, RONALD BROWN AND DONALD HAKE

Southern Illinois University

The major purpose of this study was to determine relationships between the 1956 edition of the Progressive Matrices and standard IQ tests, using U. S. children as subjects. Raven [5] has reported an *r* of .90 between the Matrices and the Stanford-Binet for English children. A secondary purpose was to determine whether the Matrices would differentiate between institutionalized and non-institutionalized retardates and between both these groups and normal children.

The sample consisted of 104 retarded children (56 institutionalized and 48 enrolled in educable mentally handicapped classes) and 67 normal children enrolled in the same schools as the EMH children.[2] The age range was 8-13 years with mean CA's of the retardates and normals being eleven years eight months and eleven years five months, respectively. There were no significant differences in mean CA between any of the sex-groupings of either the retarded or normal groups.

The retarded children were tested with the Wechsler Intelligence Scale for Children. The mean Full Scale IQ for all retardates was 65.29 (SD = 8.64), with the institutionalized group having slightly lower IQs than the EMH children (mean IQ = 62.71 and 68.06, respectively, with SD's of 7.60 and 8.80). The normal children were tested with the California Test of Mental Maturity and their mean Total Scale IQ was 110.73 (SD = 15.85). There were no significant IQ differences between boys and girls in any of the groups.

The Matrices were administered individually under standard conditions to all *Ss*. Raw scores were used for purposes of analysis. Partial *rs*, with CA held constant, were used to compare the mean IQ and Matrices scores for all groups. Differences in mean Matrices scores between groups were determined by means of the *t*-test for independent samples.

Results

Results of the correlational analysis revealed partial *rs* of .506 (all retarded *Ss*), .508 (institutionalized retardates), and .511 (EMH retardates) between the WISC and Matrices scores. A slight sex difference in favor of girls' scores was observed (*r* = .551 girls, and .484 boys). These *rs* are comparable to those reported between the Wechsler-Bellevue Form I and the 1938 Matrices for retarded children [7]. The partial *r* between CTMM and Matrices scores for all normals was .541, with girls' *rs* again higher than those for boys (.643 and .432, respectively). These *rs* are slightly higher than those reported between the 1947 Matrices and group intelligence test scores [6]. They are somewhat lower than those reported between the 1938 Matrices and the Columbia Test of Mental Maturity [1] and considerably lower than those reported by Raven between the 1956 Matrices and the Stanford-Binet [5].

Highly significant differences were found when the Matrices scores of all retarded and normal children were compared. Mean scores of 17.55 and 27.95, respectively, with SDs of 4.44 and 5.47, yielded a *t*-ratio of 13.641, significant at the .001 level. Comparable *t*-ratios were found when retarded and normal boys' scores and girls' scores were compared (*t* = 12.387 and 9.426, respectively). However, although the institutional retardates had slightly lower mean Matrices scores than the EMH group, no significant differences were found. There were no sex differences within either of the retarded groups or the normal group. These results suggest that the 1956 revision of the Progressive Matrices differentiates retarded and normal children practically as well as IQ scores and that these Matrices demonstrate difference-trends between institutionalized and non-institutionalized retardates of the same order as IQ scores.

Because of the between-group differentiations just noted, the significant relationships between these Matrices and IQ scores already given, and the economy in administering and scoring which the Matrices offer, it would appear that the 1956 edition of the Progressive Matrices offers possibilities as a screening device for U. S. children, both retarded and normal, in the age range covered by this study.

References

1. Barratt, Ernest. Relation of the PM (1939) and CMMS to the WISC. *J. consult. Psychol.*, 1956, *29*, 294-96.
2. Bolton, F. B. Excerpts on Raven's Progressive Matrices (1938). *J. Educ. Res.*, 1955, *48*, 629-33.
3. Green, M. W. and Ernest, J. C. Normative data on the PM (1947). *J. consult. Psychol.*, 1955, *19*, 139-42.
4. Martin, H. W. and Wiechers, S. E. Raven's Progressive Matrices and the WISC. *J. consult. Psychol.*, 1954, *18*, 143-44.
5. Raven, J. C. *Guide to Using the Coloured Progressive Matrices* (1956). London: H. K. Lewis & Company, Ltd., pp. 40.
6. Rimoldi, H. C. A note on Raven's Progressive Matrices Test. *Educ. Psychol. Meas.*, 1948, *8*, 347-52.
7. Stacey, C. and Carleton, F. The relationship between Raven's Coloured Progressive Matrices and two tests of general intelligence. *J. clin. Psychol.*, 1955, *11*, 84-85.

[1]This study was conducted as part of a research project sponsored by the U. S. Office of Education Cooperative Research Program, Contract No. 176(6471).

[2]Acknowledgement is hereby made for cooperation given by Joseph Albaum, M. D., Superintendent, and Mr. William Chambers, Chief Psychologist, Lincoln (Illinois) State School and Colony, and by principals and teachers in Carbondale, Christopher, Herrin, Marion, and Murphysboro, Illinois.

A COMPARISON OF WAIS SCORES OF HEMIPLEGIC PATIENTS WITH AND WITHOUT APHASIA[1]

SAMUEL HIRSCHENFANG

Kings County Hospital Center, Brooklyn, New York

Problem

This study investigates the possibility that left hemiplegic patients without aphasia do poorer on the Wechsler Adult Intelligence Scale [4] than patients having right hemiplegia with aphasia. Morrow and Mark [3] found that patients having lesions in the left hemisphere, showed greatest defect and greater loss in both verbal and performance IQ. Wechsler [5] found that in some cases of left cortical damage with aphasia, the patient may do much better on performance tests than the patient with right cortical damage and no aphasia symptoms. Wechsler's results agree with Weisenburg and McBride [6] who showed that right hemisphere damage results in lowered performance tests. Karlin, Hirschenfang, Miller and Rich, [2] similarly found that left hemiplegic patients show considerably more disturbances in performance functioning than right hemiplegics with aphasia. Karlin and Hirschenfang [1] found that both right and left hemiplegic patients showed disturbance on the performance area of the WAIS, and that the left hemiplegics demonstrated more disturbance in visuospatial discrimination, in spite of little visual agnosia and no language dysfunction as indicated by no aphasia.

Method

The Wechsler Adult Intelligence Scale (WAIS) was administered to two groups of 25 institutionalized hemiplegic patients each: 25 right hemiplegics with and without aphasia, and 25 left hemiplegics without aphasia. The subjects were matched as closely as possible for chronological age, and number of years of formal academic training.

The mean age of the right hemiplegics was 56 years and for left hemiplegics, 60 years. The age range of the two groups was 22-78 years for right hemiplegic and 46-83 years for left hemiplegics.

Results

Table 1 reveals that both groups of hemiplegics with and without aphasia did poorly on the performance scale of the WAIS. The Digit Symbol subtest involving visual acuity, motor coordination and speed, was found to be most difficult by both groups. In the right hemiplegic patients, this might be explained by the fact that all the patients were previously right handed, and as a result of the stroke, were unable to use the preferred right arm. However, all the left hemiplegic patients were using their previously dominant arm and still had difficulty in these tasks. In comparing still further the performance tests of the WAIS, the left hemiplegic patients without aphasia showed lower scores than the right hemiplegic patients, with and without aphasia. The two verbal subtests of the WAIS, found to be most difficult by both groups, were Digits Backward and Similarities. These results would agree with the psychometric test patterns for organic brain pathology as described by Wechsler [5].

[1]From the Department of Physical Medicine and Rehabilitation and the Department of Neurology, Jewish Chronic Disease Hospital, Brooklyn, New York. This investigation is supported by a Research Grant #B1382 from the National Institutes of Health, Public Health Service, Principal Investigator—Isaac W. Karlin, M. D.

Table 1. Mean Scores of WAIS Subtests in Two Groups of Hemiplegic Patients

WAIS Scales	Right Hemiplegia with and without aphasia N = 12*	Left Hemiplegia, without aphasia N = 25
Verbal		
Information	6.9	8.1
Comprehension	6.0	7.6
Arithmetic	7.5	6.8
Similarities	4.0	5.0
Digit Span	5.0	7.2
Digits Forwards**	4.8	5.9
Digits Backwards**	2.5	3.1
Vocabulary	6.0	8.1
Performance	N = 25	N = 25
Digit Symbol	1.4	1.4
Picture Completion	4.2	3.1
Block Design	3.6	2.2
Picture Arrangement	3.3	1.5
Object Assembly	2.9	2.0

*13 right hemiplegic patients could not be tested, due to severe expressive aphasia.

**Scores for Digits Forwards and Backwards are *raw* scores.

SUMMARY

This study investigated whether left hemiplegic patients without aphasia demonstrate more difficulty on a test of intelligence than do right hemiplegic patients, with and without aphasia. The WAIS was individually administered to 25 left hemiplegics and 25 right hemiplegics. The left hemiplegics without aphasia show consistently poorer performance scores than the right hemiplegic patients, with and without aphasia. It is suggested that, on more careful examination, the left hemiplegics might present greater problems in rehabilitation, than the right hemiplegics with and without aphasia.

REFERENCES

1. KARLIN, D. and HIRSCHENFANG, S. A comparison of visuosensory and visuomotor disturbances in right and left hemiplegics. Accepted for publication in the *Amer. J. Ophthalmology*.
2. KARLIN, I. W., HIRSCHENFANG, S., MILLER, M. H., and RICH, H. A comparative evaluation of right and left hemiplegic patients, with and without aphasia. Submitted for publication.
3. MORROW, R. S., and MARK, J. C. Correlation of intelligence and neurological findings on 22 patients autopsied for brain damage. *J. consult. Psychol.*, 1955, *19*, 283-289.
4. WECHSLER, D. *Wechsler Adult Intelligence Scale*. New York: The Psychological Corporation, 1955.
5. WECHSLER, D. *The Measurement and Appraisal of Adult Intelligence* (4th Edition). Baltimore: The William and Wilkins Co., 1958.
6. WEISENBURG, T., and MCBRIDE, K. *Aphasia*. New York: Commonwealth Fund, 1935.

EFFECT OF MENTAL ILLNESS ON INTELLIGENCE TEST SCORES

RICHARD M. GRIFFITH AND VIVIAN H. TAYLOR

Veterans Administration Hospital, Lexington, Kentucky, and the University of Kentucky

If mental illness impairs performance, as is generally assumed, the correlation between intelligence test scores and grade level of mental patients should be lower than for the normal. The files of records accumulated through the years by the Psychology Service in a large NP hospital yielded a sample of 356 containing both intelligence test results and personal data. The tests were restricted to the various Wechsler versions (I, II, WAIS). The grade level noted on the record was accepted as educational attainment. Diagnoses were ignored, the cases being taken as they came, with the exception that "organics" were handled separately.

The product-moment correlation between grade level and IQ was .62. The 41 records of organics gave a correlation of .34, a value significantly lower than that for the other NP patients ($p < .05$, two-tailed test). Miner [1, p. 67], reviewing the more recent studies correlating education and intelligence, reports a correlation of .64 in the original Wechsler standardization sample which dropped to .53 when mental defectives were omitted; another study found a comparable value of .56. The size of the present correlation is more impressive because mental defectives were supposedly screened out of inductees before service, although some veterans do score at the defective level in the hospital.

The fact that the correlation between IQ and education is very nearly as high with hospitalized patients as in the normal population tends to contradict the assumption that mental illness markedly distorts intelligence test scores. That the correlation is considerably lower for organics, whose performance does suffer, more or less validates the conclusion.

A large number of the diagnostic protocols surveyed in the present study asserted on the basis of one bit of evidence or another that the patient's intelligence was probably higher than the measured value. It would seem that very often this qualification was ill-founded.

REFERENCES

1. MINER, J. B. *Intelligence in the United States*. New York: Springer Publishing Co., 1956.

EDITORIAL OPINION (continued from page 353)

who are unable to procure the Ph.D. under the present research-oriented doctoral program, would it not be better to establish some mechanism whereby they could receive the highest levels of training in their specific fields without requiring an original research project?

A side issue relating to this controversy is the question of whether the APA should place its official imprint on the transactions of component committees and conferences until their content is reviewed by the APA Council. This particular instance carries more insidious implications than most similar publications and it would seem desirable to resolve the issues by early submission to the entire membership.

F. C. T.

EDITORIAL OPINION

GRADUATE EDUCATION IN PSYCHOLOGY

Res ipsa loquitur. This basic legal concept means "let an action (or transaction) speak for itself." In this connection, let us examine the published intent and conclusions of the APA-sponsored Miami conference on "Graduate Education in Psychology" in contrast with the implied intent underlying its organization and its interpretation by the general public. Although the conference's message is expanded into a 97 page monograph quoting the standard cliches and reassurances traditionally rationalizing such gatherings in terms of the accepted credos of time and place, its essential conclusion is contained in the following: ". . . . this Conference has gone firmly on record that it approves the development of nondoctoral training for some who are to teach psychology, and for technical specialties, some of which may be psychological in nature."[1] No amount of weasel wording intended to reassure our doubts and fears that such a recommendation essentially would operate to lower the standards and competence for clinical practice can alter the fact that the meaning of such an action speaks louder than words. At a time when the long struggle of the APA to establish the Ph.D. degree followed by appropriate in-service training as the minimum level of competence suitable for clinical practice has just begun to win public acceptance, we now find a responsible APA group most emphatically supporting lower levels of competence.

We are unimpressed by such arguments as that sub-doctorally trained personnel will always exist so we might as well legalize their status. Or that the current public demand for psychological services is so great that it can only be met by sub-doctoral personnel. Or that such levels of training could be made adequate by standardization and improvement under APA supervision. Or that the position of professional psychology would be preserved by labelling subdoctoral personnel by other technical terms. Or that lower levels of personnel really do not need doctoral training. Conversely, we might argue that it is better not to try to keep up to public demand by supplying incompetent personnel. And that a superior objective would be to enlarge training facilities to provide the needed supply of the best trained personnel that can be produced at time and place. We must insist that clinical psychology is such a complex subject that no valid shortcuts to competence can ever exist apart from the most intensive long-term training. It is a tactical error in long term planning to accept inferior short-term makeshift solutions to problems which can only be solved properly by the most diligent effort of all concerned.

This controversy comes at a time when professional psychology is struggling towards improved interprofessional relations, especially with medicine. Our newfound hardwon professional status is still so tentative that nothing must be allowed to sabotage the high standards upon which it is based. Professional psychology (represented by the APA) cannot tolerate any tampering with the doctoral standard of training which represents an absolute minimum. To be certain, there will always exist personnel with subdoctoral levels of training and substandard competence but they must not be allowed to designate themselves as in any way certified or approved by official organizations. To give official professional recognition to substandard personnel at this time would sabotage the status of professional psychology and our interprofessional relations.

Rather than to support any weakening of professional standards by letting down the bars to all sorts of subdoctoral technicians, why not embark upon a crash program to raise all personnel to doctoral levels? Admitting that there are a large number of subdoctoral personnel who are performing competently in the field but

---

[1] ROE, ANNE *et al* (Eds.) *Graduate Education in Psychology.* Washington: American Psychological Association, 1959, pp. 97.

(continued on page 352)

Wiley BOOKS

coming in August . . .

HANDBOOK OF RESEARCH METHODS IN CHILD DEVELOPMENT

Edited by Paul Henry Mussen, Associate Professor of Psychology, *University of California, Berkeley*. The quantity and quality of research wisdom incorporated in this book is not found anywhere else in the literature of child development, or for that matter, in any other field of psychology. Each chapter has been written by an outstanding worker and teacher in the field of Child Development.

Describes and evaluates the most widely used techniques

The handbook makes available in a single source concrete descriptions and evaluations of the most widely used research techniques in many aspects of the study of child growth and behavior. In addition, it communicates, wherever possible, some of the methodological sophistication and "wisdom" that investigators gradually acquire but rarely make explicit in their publications. The book includes only the methods that have actually been adopted with children and have not been thoroughly described in other volumes. Discussions are limited to the research—as opposed to diagnositc or therapeutic—uses of the methods.

A valuable reference for both student and investigator

In a sense, *Handbook of Research Methods in Child Development* is addressed directly to present and potential investigators who can use it as a basic reference and a guide to research matters outside their own research specialities. It should also be valuable to advanced scholars and graduate students interested in a balanced picture of the techniques applied in child study. Although most of the chapters deal primarily with research in child development, they are of broad interest and would be useful as source material in courses devoted to research methodology in psychology and other behavioral sciences.

To facilitate reference, the book is divided into five main sections of specialization. Part I is concerned with general research design and the use of observation and experimental methods in research. The next section consists of a study of biological growth and development. Part III examines the cognitive processes (perception, learning, language, etc.). Part IV deals with the major research tools and areas of study in the fields of personality development and social behavior, and the last section is devoted to the study of the child's social environment.

1960. Approx. 1056 pages. Prob. $15.50.

Send for an examination copy.

JOHN WILEY & SONS, Inc.

440 Park Avenue South, New York 16, N. Y.

Announcing New Services

RESEARCH DESIGN ASSOCIATES

(A Division of the *Journal of Clinical Psychology*)

The Editorial Staff of the *Journal of Clinical Psychology* is now organized to provide professional consultation services on all aspects of research design, statistical evaluation, data processing, editorial criticism and publication problems. Specialized consultants now are available in the areas listed below.

PROJECT ORIGINATION	RESEARCH PLANNING
EXPERIMENTAL DESIGN	DATA PROCESSING
STATISTICAL DESIGN	STATISTICAL EVALUATION
CRITICAL EVALUATION	DATA ANALYSIS
EDITORIAL CRITICISM	PUBLICATION PROBLEMS

These services are offered in connection with research problems in any of the behavior sciences. Our consultants are particularly competent in relation to clinical problems in clinical psychology, psychiatry and the applied social sciences.

These services will be available at all stages of research design, data collection and data evaluation. Please consult us if you desire an independent evaluation of research designs or work in progress.

Consultations may be arranged either by mail or by a direct visit to your installation. Where indicated, several consultants will render opinions concerning various aspects of the project. Gratis estimates of costs may be obtained by sending an outline of the problems involved together with a statement of the services desired. For further information please write to the address given below.

RESEARCH DESIGN ASSOCIATES
JOURNAL OF CLINICAL PSYCHOLOGY
5 Pearl Street
Brandon, Vermont

"The new and different in tests. . . ."

Could you use objectively obtained estimates of strengths and interrelationships of the id, ego, and superego in your diagnostic work? In your research on personality? Are you just plain curious about how these basic concepts of psychoanalytic theory might be quantified?

The

I

E

S

TEST

utilizes four independent subtests to measure important aspects of id, ego, and superego functioning. Materials are easy to handle. Tasks are disarming and are intrinsically interesting to almost anyone. Evaluation takes as much time, thought, and sophistication as for any other test giving important basic information about the personality.

•

Obtain information and/or order (complete set $28.50) from

PSYCHOLOGICAL TEST
SPECIALISTS
Box 1441
Missoula, Montana

"The new and different in tests. . . ."

SEX CRIMES *by* **John Drzazga,** *Police Dept. N. Y. C. (Retired).* A unique and extensive study of the "modus operandi" of professional purveyors of vice and the laws restricting such conduct. Although dealing principally with legal aspects of sexual conduct, Mr. Drzazga also covers sexual history and customs, techniques used by perverts, types of sexual deviations or practices, and commercial vice. Based on exhaustive research, much of the material APPEARS IN PRINT HERE FOR THE FIRST TIME. *(Police Science Series edited by V. A. Leonard). Pub. date May '60*

AN APPLICATION OF PSYCHOANALYSIS TO EDUCATION *by* **Richard M. Jones,** *Brandeis Univ., Waltham, Mass.* A detailed narrative of tape-recorded verbatim records of classroom exchanges of an experimental high school senior group over a school year. An intricate pattern of findings on significant changes in projectivity, intraception, attitudes of discrimination, and prejudiced beliefs forms a graphic base for the author's conclusions and further research suggestions. *(Amer. Lec. Psychology edited by Molly Harrower). Pub. date May '60, 136 pp., $5.50.*

PSYCHOANALYTIC CONCEPTS OF DEPRESSION *by* **Myer Mendelson,** *Univ. of Pa. School of Med., Philadelphia.* Dr. Mendelson's book represents the FIRST SURVEY comprehensive enough to place recent developments in the theory of depression in proper perspective. Attention is focused on constitutional and psychological factors predisposing to depression; the relation of a central anxiety situation in human development to depression, the role of obsessional elements; aggression; orality and introjective mechanisms; depressive reactions; and treatment. *Pub. date June '60*

CHEMICAL SENSES IN HEALTH AND DISEASE *by* **H. Kalmus** *and* **J. S. Hubbard,** *both of Univ. College London.* Why are the cook and the perfumer still artists and not scientists? Why do we nevertheless encounter very solid, practical success in the large scale manufacture of food, drinks, flavoring, cosmetics, and deodorants? From University College London have come some very interesting answers to these questions. After introducing general aspects of sensory physiology, the authors consider the role, physiology, and psychology of taste and smell. *Pub. date June '60, 104 pp., 10 il., (Amer. Lec. Living Chemistry edited by I. Newton Kugelmass), $3.75*

7 NEW BOOKS

of special interest to You!

BABY TALK *by* **Morris Val Jones,** *School for Cerebral Palsied Children, San Francisco, Calif.* Most parents struggle with the decision to seek professional help; many try to determine the causes for their children's speech deviations; and nearly all of them ask what they can do at home. BABY TALK was written just for such parents. They are given "common sense" advice about their role in helping children overcome this speech problem and develop normal articulation. A book that any physician or psychologist can recommend with complete confidence. *Pub. date June '60, 104 pp., 9 il., $4.50*

CREATIVE VARIATIONS IN THE PROJECTIVE TECHNIQUES *by* **Molly Harrower,** *Temple Univ. Med. Center;* **Pauline Vorhaus,** *Veterans Administration, N. Y. C.;* **Melvin Roman** *and* **Gerald Bauman,** *both of Albert Einstein Med. College.* Clinical psychologists with analytic experience will welcome these new and practical techniques — PROJECTIVE COUNSELING, THE STRUCTURED INTERVIEW, and INTERACTION TESTING. ". . . this book will start clinicians to develop confidence in their technique and give them a variety of tools to make the logical step from diagnosis to therapy."—From the Foreword by Bruno Klopfer. *Pub. date May '60, 160 pp., 32 il., (Amer. Lec. Psychology edited by Molly Harrower), $8.50*

EVALUATION OF CHANGES ASSOCIATED WITH PSYCHIATRIC TREATMENT *by* **Marvin Reznikoff** *and* **Laura C. Toomey,** *both of the Institute of Living, Hartford, Connecticut.* What REALLY happens to a patient receiving psychiatric treatment? HERE IS YOUR ANSWER—an overview of current thinking and achievements in the area of clinical research involving the treatment of psychiatric patients. Includes fascinating research on all types of treatment—psychotherapy, shock therapies (insulin, electric, metrasol), lobotomy, and tranquilizing medication. *Pub. date June '59, 132 pp., $4.50*

Be on the lookout for another important new book . . .

KRAFFT-EBING'S ABERRATIONS IN SEXUAL LIFE by Alexander Hartwich

---

CHARLES C THOMAS - PUBLISHER

301-327 East Lawrence Avenue

Springfield, Illinois

DEVEREUX SERVES

A FULL EDUCATIONAL PROGRAM, extending from kindergarten through senior high school and beyond, is offered at Devereux for emotionally disturbed, as well as slow-learning children.

Instruction is highly individualized and planned to take advantage of the specialized learning techniques which are most effective for these children.

The curricula of the Devereux units include a solid academic program for college preparatory work; a vocational and commercial program; and programs of artisanship or craftsmanship, stressing the use of special tools and skills, either as a preparation for admission to a vocational or trade school or as a means of attaining self-confidence and personality development.

CLINICAL STAFF

J. Clifford Scott, M.D.
Edwin H. Abrahamsen, M.D.
Aurelio Buonanno, M.D.
Charles M. Campbell, Jr., M.D.
Fred J. Culeman, M.D.
Ruth E. Duffy, M.D.
William F. Haines, M.D.
Robert L. Hunt, M.D.
Richard H. Lambert, M.D.
Leonardo Megran, M.D.
Joseph J. Peters, M.D.
Alvis J. Scull, M.D.
Jacob S. Sherson, M.D.
Albert S. Terzian, M.D.
Walter M. Uhler, M.D.
Tirso L. Vinuesa, M.D.

Lance Wright, M.D.
F. Ellsworth Henry, S.T.D.
Milton Brutten, Ph.D.
William J. Cohen, Ph.D.
Dorothy E. Conrad, Ph.D.
Sidney L. Copel, Ed.D.
Michael B. Dunn, Ph.D.
Shirley M. Jahnson, Ph.D.
Noel Jenkin, Ph.D.
John R. Kleiser, Ph.D.
Murray Levine, Ph.D.
Henry Platt, Ph.D.
Edgar A. Smith, Ed.D.
George Spivack, Ph.D.
Herbert A. Sprigle, Ph.D.
Anne Howe, M.S.

Kenneth E. Evans, B.S.

Psychoanalytic Consultants

G. Henry Katz, M.D.
Herbert H. Herskovitz, M. D.

THE DEVEREUX FOUNDATION

A nonprofit organization — *Founded 1912*

Devon, Pennsylvania

Santa Barbara, California — *Victoria, Texas*

SCHOOLS
COMMUNITIES
CAMPS
TRAINING
RESEARCH

HELENA T. DEVEREUX
Administrative Consultant

EDWARD L. FRENCH, Ph.D.
Director

WILLIAM B. LOEB
Treasurer

Professional inquiries for Eastern Schools should be directed to CHARLES J. FOWLER, *Registrar*, Devereux Schools, Devon, Pa.; for Pacific Coast Schools, to KEITH A. SEATON, *Registrar*, Devereux Schools in California, Santa Barbara, Calif.; Southwestern residents address Devereux Schools in Texas, Box 336, Victoria, Tex.